Opaque Interaction of Internal Merge and Agree

Doreen Georgi*

Abstract

In this paper I present an argument for a strictly derivational model of syntax based on timing of operations. Empirical evidence comes from opaque interactions of elementary operations which show that internal Merge (IM) must be split into two types: IM to intermediate landing sites and IM to final landing sites. The split is motivated by the following observation: When both types of IM are triggered by the same head H, they apply at different points in the derivation. This becomes visible once they interact with Agree: In some languages, IM to final landing sites feeds/bleeds Agree initiated by H, whereas IM to intermediate landing sites has the opposite effect, i.e., it counter-feeds/counter-bleeds Agree. This effect can be derived by ordering operation-inducing features on H: One type of IM applies before and the other after Agree. The general implication is that Agree not only needs to be ordered with respect to Merge; a more fine-grained approach is needed that distinguishes between different types of (internal) Merge. Furthermore, reordering of the operation-inducing features on H predicts a certain range of cross-linguistic variation. Based on the attested variation, I argue for the need of extrinsic ordering.

1. Introduction

In this paper I present an argument for a strictly derivational model of syntax based on the timing of elementary syntactic operations. The evidence comes from opacity effects that show that internal Merge (IM) is not a uniform operation; rather, IM must be split into two types: (i) IM to intermediate landing

*For valuable comments on earlier versions of this paper I would like to thank the participants of the syntax colloquium at the University of Leipzig as well as the audiences at ConSOLE XXI (Potsdam, January 2013) and at the workshop Interaction of Syntactic Primitives (35th Annual Conference of the German Linguistic Society, March 2013). I profited enormously from the discussions on rule ordering with Anke Assmann, Fabian Heck, Gereon Müller, Martin Salzmann and Philipp Weisser. This research is supported by the DFG (Deutsche Forschungsgemeinschaft) within the project "Lokale Modellierung nicht-lokaler Abhängigkeiten in der Syntax" (Local modeling of non-local dependencies in syntax): http://www.uni-leipzig.de/~lomo.

Rule Interaction in Grammar, 413–461
Fabian Heck & Anke Assmann (eds.)
Linguistische Arbeits Berichte 90, Universität Leipzig 2013

sites, triggered by edge features, and (ii) IM to final landing sites in a movement chain, triggered by other features such as a wh-feature on C, the EPP on T, etc. This split is empirically motivated by the observation that when both types of IM are triggered by the same head H, they apply at different points in the derivation. This effect becomes visible once they interact with Agree: In some languages, IM to final landing sites feeds/bleeds Agree relations initiated by H, whereas IM to intermediate movement sites has exactly the opposite effect, i.e., it counter-feeds/counter-bleeds Agree. In the latter cases, the interaction of IM and Agree is opaque. The term opacity characterizes rule interactions that are non-transparent: When looking at the output of an opaque interaction, it is unclear (a) why a certain operation has not applied although its context is given (counter-feeding) or (b) why an operation has applied although its context is not given (counter-bleeding), cf. Kiparsky (1976). The cases at hand are opaque because internally merged XPs land in the same position SpecH (the Spec of the head H that triggered IM) whether IM is driven by edge features (intermediate IM) or by other features (final IM). The output structure is identical; nevertheless, the two types of IM have different consequences for Agree initiated by H, which is reflected in the presence or absence of a morphological marker. The effect can be modeled by ordering operation-inducing features on the triggering head H: edge feature-driven IM applies *after* Agree and non-edge feature-driven IM applies *before* Agree initiated by H. The consequence of this order is that the former type of IM applies *too late* to change possible Agree relations (the DP that is to be internally merged is still in its base position when Agree applies); the latter type of IM changes structural relations *before* Agree applies and can thus feed or bleed Agree relations (depending on the input), because Agree is structure-sensitive. This analysis of opacity effects crucially relies on timing of elementary operations and thereby provides an argument for a strictly derivational syntax (cf. also Řezač 2004, Heck and Müller 2007).

I will show that opacity effects with this abstract pattern can be found with a number of functional heads in the clause on the basis of the phenomena that follow: the anti-agreement effect in Berber, defective intervention in Romance and Icelandic, and possessor case/agreement in Hungarian. The analysis presupposes that intermediate movement steps are triggered by designated features (edge features) that are different from the triggers of final movement steps (contra Abels 2012).

Furthermore, I will argue on the basis of cross-linguistic variation that not all of the attested orderings of operation-inducing features on a head H are predicted by principles of the grammar (contra Pullum 1979); in particular, the Cyclic Principle (even in its strongest form as formulated in McCawley 1984, 1988) has nothing to say about orderings of operations that are triggered by the same head. Extrinsic ordering (or: parochial ordering in Pullum's 1979 terms) is thus necessary after all to account for the variation. Interestingly, not all logically possible orderings of the triggers of the two IM types and Agree seem to be attested for the phenomena examined in this paper. I suggest an account of this asymmetry which is based on specificity-driven ordering of operation-inducing features on a head. In addition, the variation found with defective intervention effects provides evidence that wh-movement uses SpecT as an intermediate landing site (for a recent defense of the opposite view see Abels 2003, 2012 and references cited there).

The paper is structured as follows: Section 2 gives a short overview of the types of rule interactions and the operations that interact in minimalist syntax. Section 3 introduces the theoretical assumptions and illustrates the abstract patterns of interactions that will be encountered in the data. Furthermore, it is shown how the opacity is resolved under the given assumptions. Section 4 presents concrete instantiations of the abstract patterns on the basis of several phenomena and presents detailed derivations. Section 5 discusses consequences of the analysis and examines whether its predictions with respect to cross-linguistic variation are borne out. Furthermore, the implications for the extrinsic/intrinsic dichotomy are discussed. Finally, section 6 concludes.

2. Rule Interaction in Grammar

This section introduces the four basic types of rule interactions: feeding, bleeding, counter-feeding and counter-bleeding. Furthermore, I present the syntactic primitives that can interact in Minimalism.

2.1. Types of Rule Interactions

In early transformational grammar, the grammar consists of two core components: the base component which generates the underlying structure (deep structure) of a linguistic expression, and a transformational component which relates the underlying structure to the surface structure of a linguistic expression (cf. Chomsky 1957, 1965 on syntax and Chomsky and Halle 1968 for application of this model to phonology). The transformational component consists of a number of rules that apply to the underlying structure and map it onto another structure. Often, more than one rule applies to derive the surface structure of a linguistic expression from its deep structure. It has been observed that in this case, the rules may interact in intricate ways: The application of a rule R_1 may facilitate or block the application of another rule R_2. The former case is an instance of feeding, the latter is an instance of bleeding. Kiparsky (1968, 1971, 1976) divided rule interactions into two types: transparent and opaque interactions. Transparent interactions comprise feeding and bleeding; opaque interactions comprise counter-feeding and counter-bleeding. These terms are defined as follows (X '>' Y means that X applies before Y):

(1) *Transparent interactions*:

a. Feeding:
   (i) A rule R_1 creates the context for the application of a rule R_2.
   (ii) *Example* R_1: A → B > R_2: B → C

b. Bleeding:
   (i) A rule R_1 destroys the context for the application of a rule R_2.
   (ii) *Example* R_1: A → B > R_2: A → C

(2) *Opaque interactions*:

a. Counter-feeding:
   (i) A rule R_1 creates the context for the application of a rule R_2 and should thus feed R_2.
   (ii) However, empirical evidence shows that R_2 has not applied although R_1 has.
   (iii) A rule has not applied although its context is given.
   (iv) *Example* R_2: B → C > R_1: A → B

b. Counter-bleeding:
   (i) A rule R_1 destroys the context for the application of a rule R_2 and should thus bleed R_2.
   (ii) However, empirical evidence shows that R_2 has applied although R_1 has as well.
   (iii) A rule has applied although its context is not given.
   (iv) *Example* R_2: A → B > R_1: A → C

The opaque interactions can be characterized as follows: It is expected that a rule R_1 facilitates (feeds) or blocks (bleeds) the application of another rule R_2, but this is not the case. On the surface, (i) rule R_2 has not applied although its context is created by the prior application of R_1 (counter-feeding), or (ii) R_2 has applied although its context should have been destroyed by the prior application of R_1 (counter-bleeding). Opacity can be described by reversing the order of operations: If R_1 applies before R_2 the former should feed or bleed the latter. But since this is not the case, R_2 must have applied before R_1; R_1 applied too late to facilitate or block the application of R_2.

Rule interactions have been a major topic in phonology and syntax since the earliest days of generative grammar. Opacity was first described by Chomsky (1951) for Hebrew phonology. For an overview of interactions of phonological rules see e.g. Chomsky and Halle (1968), Anderson (1969, 1974), Koutsoudas et al. (1974), Kenstowicz and Kisseberth (1977, 1979), Baković (2011); on rule interaction in morphology see Embick (2010), Arregi and Nevins (2012). As for syntax, the most comprehensive treatment of this topic is Pullum (1979); see also Ross (1967), Williams (1974), Kayne (1975), Perlmutter and Soames (1979), McCawley (1984, 1988), Řezač (2004), Lasnik (2001) and Brody (2002).

The topic of this paper are opaque interactions in syntax. A well-known example for counter-bleeding in syntax that has been recurring in the transformational literature is the interaction of Reflexivization and Imperative Subject Deletion (ISD) in English (cf. McCawley 1988: ch.6 and Pullum 1979: ch.1).

(3) *Informal definitions of transformations*:
  a. Reflexivization: If the direct object of a transitive verb is coreferent with the local subject, it must be realized as a reflexive pronoun.
  b. Imperative subject deletion: A deep structure 2nd person subject is deleted when it is the subject of an imperative.

In imperatives, the direct object is reflexivized if it is 2nd person and thus coreferent with the (phonologically empty) subject:

(4) *Reflexivization in imperatives in English*
  a. Defend yourself!
  b. *Defend you!

On the surface, it is unclear why Reflexivization could apply in (4-a) given that its context is not given because there is no coreferent subject. Since Reflexivization makes reference to the subject of a transitive verb, deletion of the subject by ISD could bleed Reflexivization, but it does not (counter-bleeding). This implies that Reflexivization applies before ISD, at a point in the derivation where the subject is still present; it is deleted only afterwards.

2.2. Rule Ordering in Minimalism: Conflicts in the Derivation

Minimalism is a derivational approach to syntax (cf. Chomsky 1995 et seq.). Syntactic structures unfold incrementally in a bottom-up fashion by successive applications of the elementary operations Merge and Agree (Chomsky 2000, 2001). These are defined as follows:

(5) *Merge and Agree*:

a. Merge (external and internal) is a structure-building operation. Merge is triggered by structure-building features [•F•].

b. Agree relates functional heads and arguments. It is (among other things) responsible for argument encoding: (i) ϕ-features are copied from DPs onto functional heads and (ii) case values are assigned by functional heads to DPs.
Agree applies under c-command and is triggered by probe features [∗F∗].

It is these basic operations that can interact in Minimalism. In particular, the structure-sensitive Agree operation can be fed or bled by Merge which builds structures and which thus may change structural relations.

That probe features and structure-building features must be ordered has been motivated independently: Some functional heads trigger more than one operation. Little v, for example, triggers ϕ-Agree and (external) Merge of the external argument. It thus has the following features: v { [•D•], [∗ϕ∗] }. The T head in English is another example: It triggers Agree with the subject DP and (internal) Merge to SpecT (the EPP property); it thus bears the following features: T { [•D•], [∗ϕ∗] }. Hence, at the point of the derivation where v or T is merged, a conflict arises (Müller 2009): The relevant head could either first discharge its structure-building feature and then the probe feature, or it could proceed the other way around. Assuming that only a single operation can apply at any given stage of the derivation, these operation-inducing features on v and T must be ordered.

It has recently been argued, although often rather implicitly, that the order of features on a head that triggers more than one operation is not arbitrary. Van Koppen (2005), Béjar and Řezáč (2009), Halpert (2012), and Assmann and Heck (2013) argue for a strict order of the two operations on certain heads.[1] Müller (2009), Heck and Müller (2007), Lahne (2008*a*), and Assmann, Georgi et al. (2013) on the other hand, argue that the order of Merge and Agree on v and T is in principle free and may vary between languages, i.e. it is determined in a language-specific fashion. The choice of the order in a given language is responsible for cross-linguistic variation, e.g. for the pattern of argument encoding (ergative vs. accusative), the extractability of core arguments and word order variation.

A simple example for a fixed order of Merge and Agree on T comes from English subject-verb-agreement where the operations are in a counter-bleeding relation: The T head is standardly assumed to trigger ϕ-Agree with the closest DP in its c-command domain, which is the subject DP base-generated in Specv, cf. (6). In addition, the T head has a [•D•]-feature (the EPP-feature) that triggers internal Merge of the subject to SpecT.

(6) *Basic structure of T′ in English*:
$[_{\text{T}'}\ \text{T}_{\{[*\phi:\square*],[\bullet\text{D}\bullet]\}}\ [_{\text{vP}}\ \text{DP}_{ext\{[\phi:\text{2SG}]\}}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ \text{V}\ \text{DP}_{int}\ ]]]]$

In (6), the subject DP is in the c-command domain of T and T could thus Agree with it in ϕ-features; on the surface, however, the subject DP is no longer in the c-command domain of T because it has undergone EPP-driven movement to SpecT. Therefore, we expect that subject-verb-agreement is bled by EPP movement, but it is not, cf. (7) (the auxiliary *be*, which is located in T, is inserted to show that the subject has undergone EPP movement to SpecT).

[1] Van Koppen (2005) argues for internal Merge before Agree on the C head in languages with complementizer agreement to derive a bleeding effect; Halpert (2012) argues for the same order on an abstract head L in Zulu to account for the distribution of conjoint and disjoint morphology. Béjar and Řezáč (2009) derive hierarchy effects from the assumption that Agree between v and the internal argument of a transitive verb precedes external Merge of the external argument and may thus bleed Agree relations between v and the external argument; the proposal in Assmann and Heck (2013) also crucially relies on the order Agree before Merge on v.

(7) You are seeing John.

Hence, the operation Agree has applied although its context is not met in the output structure, cf. (8) with the subject moved to SpecT (checking the EPP on T) and the ϕ-probe on T valued by the subject.

(8) *Surface representation of TP in English*:
$[_{\text{TP}}\ \text{DP}_{ext\{[\phi:\text{2SG}]\}}\ [_{\text{T}'}\ \text{T}_{\{[\phi:\text{2SG}],\not{[}\not{\bullet}\not{D}\not{\bullet}\not{]}\}}\ [_{\text{vP}}\ \text{t}_{DP_{ext}}\ [_{\text{v}'}\ \text{v}\ [_{\text{VP}}\ \text{V}\ \text{DP}_{int}\ ]]]]]$

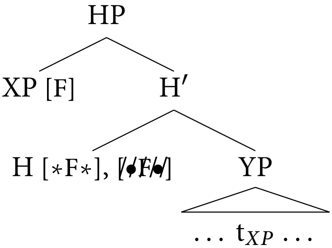
This effect is standardly derived by assuming that Agree takes place *before* the subject is EPP-moved, i.e., that the features on T are ordered as follows: T { [∗ϕ∗] ≻ [•D•] }. Agree applies when the subject DP is still in the c-command domain of T; it is EPP-moved only afterwards.

In the data that will be examined in section 4, we will see that sometimes IM to the Spec of a head H has indeed a bleeding effect with respect to the Agree relation initiated by H, suggesting that IM takes place before Agree. On the other hand, in the same language that exhibits the bleeding effect with IM, IM sometimes also behaves like EPP-movement in English in that it counter-bleeds Agree initiated by H.

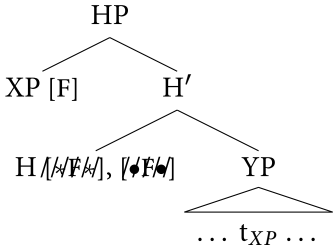
2.3. Opacity in the Present Data

The opaque interactions of internal Merge and Agree that will be presented in this paper are of the following type: A head H triggers Agree and internal Merge. It thus bears two operation-inducing features, a structure-building feature [•F•] and a probe feature [∗F∗] (a ϕ-probe or a case probe): H { [•F•], [∗F∗] }. Sometimes the XP that moves to SpecH (and thereby checks the structure-building feature on H) feeds/bleeds Agree initiated by H, and sometimes XP movement has the opposite effect in the same position, i.e., it counter-bleeds/counter-feeds Agree. Thus, IM interacts transparently and opaquely with Agree in one and the same language. This situation is abstractly depicted for (counter-)bleeding in (9) and (10) which show the surface representation of HP:

(9) *Bleeding*:



(10) *Counter-Bleeding*:



On the surface, there is an XP with the features [F] in SpecH in both (9) and (10). XP movement is triggered by a structure-building feature on H that is then discharged (indicated by crossing-out the feature). Although XP occupies the same structural position in both examples, it has different consequences for the Agree relation that H triggers as well. In (9), Agree between the probe feature [∗F∗] and the matching feature [F] on XP is bled by XP movement: The probe on H is not discharged, indicating that Agree between the probe and the XP has not taken place. In (10), on the other hand, the probe [∗F∗] on H *is* discharged, resulting in valuation of the probe in case of ϕ-Agree and in case assignment to the XP for case-Agree, respectively; hence, H and XP must have entered into an Agree relation. Comparing the representation in (10) to the one in (9), one would expect bleeding in (10) as well because the XP occupies SpecH in both examples, but there is no bleeding, viz., this is an instance of counter-bleeding.

The split between the movement types that interact transparently and opaquely with Agree is not arbitrary. The split is conditioned by the type of IM: It depends on whether XP movement to SpecH is a final or an intermediate movement step in a movement chain. The background assumption is that movement of an XP to a position where it checks a structure-building feature [•F•] (its criterial position in the sense of Rizzi 2004, 2007) does not apply in one fell swoop. Rather, XP undergoes a sequence of short movement steps until it reaches its final landing site, i.e., it moves successive-cyclically. The positions in which the XP makes a stop-over but which are not the final landing sites are intermediate landing sites. Which positions constitute intermediate landing sites for which kinds of movement is a matter of an ongoing debate. In Chomsky's (2001) phase model, Specv and SpecC are identified as intermediate landing sites for movement out of vP and CP, respectively. I will come back to this issue in section 4. But before we continue, the notion of "final" landing site needs to be clarified. The term "final" is to be understood relative to a given movement chain and not in absolute terms: It denotes the position in a chain in which an XP checks a structure-building feature (in particular, a non-edge feature, see below) that triggered its movement in the first place – the XP moves in order to check that feature. This does not necessarily mean that this position is the *ultimate* landing site for the XP, i.e., that it cannot move on from that position. An XP in the final landing site of a chain α may undergo further movement in order to check another structure-building feature in the final landing site of another chain β.[2]

What is crucial for the analysis of the opacity effect is that IM to an intermediate landing site and IM to a final landing site can be distinguished on the basis of their triggers: The former is triggered by a special kind of structure-building feature (cf. Chomsky 2000, 2001), i.e., the edge-feature [•X•], a categorially underspecified structure-building feature; the latter is triggered by other kinds of structure-building features which encode a categorial or an interpretative property of the attracted XP, e.g. the feature [•WH•] on C that triggers IM of a wh-word, the EPP-feature [•D•] on T that attracts the closest DP, or the feature [•FOC•] that attracts a focus-marked XP, etc. We will see that edge feature-driven IM counter-feeds or counter-bleeds Agree, whereas non-edge feature-driven IM feeds or bleeds Agree. The pattern that arises from this is that intermediate movement steps of XP to SpecH triggered by edge features behave as if the XP is not moved at all with respect to Agree; final movement steps to SpecH triggered by non-edge features pattern the opposite way with respect to Agree.

These effects are derived by different orderings of the two types of IM triggers relative to the Agree trigger on H: The edge feature is discharged *after* the ϕ-probe on H, whereas the non-edge feature that triggers IM is discharged *before* Agree, cf. (11). Since Agree is structure-sensitive (to c-command, see section 3 below), *early* movements (i.e., the final movement steps) can influence possible Agree relations because they change structural relations. Inter-

[2]In this respect, *final* positions as defined in this paper (and as also used in Abels 2012) are different from the *criterial* positions introduced in Rizzi (2004, 2007). Rizzi assumes that once an XP has reached a criterial position, it is frozen in that position and cannot undergo any further movement steps (= *criterial freezing*, see Rizzi and Shlonsky 2007 on consequences of this assumptions for EPP-moved XPs). For this reason, I will not use the term *criterial* in what follows, although in some instances the final position of an XP may coincide with a criterial position in Rizzi's sense.

mediate movement steps apply *too late* to have this effect, they are still in their base position when the probe on H initiates a search, just as XPs that do not undergo movement at all.

(11) a. IM to the final landing site > Agree > IM to an intermediate landing site
b. [•F•] > [$*\phi*$] > [•X•]

3. Assumptions

I assume a strictly derivational model of syntax (cf. Chomsky 1995 et seq.) in which the structure unfolds in a bottom-up fashion by successive applications of the two basic operations Merge and Agree. The basic Minimalist clause structure is given in (12). The external argument is introduced in the specifier of the functional head v. Above vP, there are two more functional projections, TP and CP.

(12) *Clause structure*:
$[_{CP}$ C $[_{TP}$ T $[_{vP}$ DP_{ext} $[_{v'}$ v $[_{VP}$ V DP_{int}]]]]]

All syntactic operations are feature-driven: Agree is triggered by probe features [*F*] and Merge is triggered by structure-building features [•F•] (I adopt the notation proposed in Sternefeld 2006, Heck and Müller 2007). Given this assumption, intermediate movement steps also need a trigger. As introduced in the discussion above, I take them to be triggered by edge features [•X•], which are categorially underspecified structure-building features. Furthermore, I assume that edge features are not freely available on functional heads; rather, they are inserted on a head H if H needs to attract an XP with a feature F although XP does not check the feature [•F•] on H (cf. *Phase Balance* in Heck and Müller 2000, 2003). For arguments that XPs must go through certain positions that are not their final landing sites see sections 4.1.3 and 4.2.3.

The operation Agree is defined in (13) (cf. Chomsky 2000, 2001). The crucial condition on Agree, which will be relevant in the derivation of opacity effects, is (13-a):

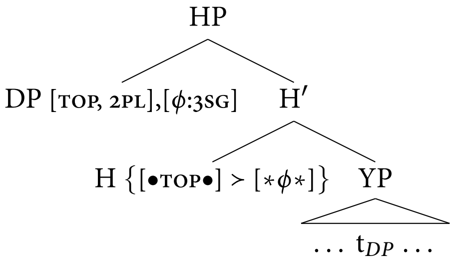
(13) Agree between a probe P and a goal G applies if
a. P c-commands G,
b. P has a feature [*F*] and G has a matching feature [F],

c. G is the closest matching goal for P.
d. Result: G values P (for ϕ-features) or P values G (for case features).

In what follows, case probes are represented as [∗c:value∗], i.e., they have a value which needs to be assigned to a DP without a value for case. If a case probe finds such a DP, it is discharged. ϕ-probes, on the other hand, do not have a value: [∗ϕ:□∗]. They seek for a value of a DP. If a ϕ-probe finds a DP with ϕ-features, the features of the DP are copied onto the probe and the probe is discharged. If Agree fails, i.e., if a probe does not find a goal, the derivation does not crash. Rather, default values are inserted either on the probe (in the case of ϕ-Agree) or on the goal (in the case of case Agree), cf. Béjar (2003), Preminger (2011). The default ϕ-features are [ϕ:3sg]; for the languages with a case split I will look at in this paper, the default case value is the nominative [c:nom]. Furthermore, I adopt the Activity Condition (Chomsky 2001): A DP that has received a case value cannot be the goal for another Agree relation, neither for ϕ- nor for case Agree. Finally, I assume that traces or copies left by movement are not visible for Agree, i.e., they cannot serve as a goal for a probe. Alternatively, one might assume that traces/copies do not exist (because movement does not leave behind anything, see e.g. Epstein and Seely 2002, Unger 2010, Müller 2011, or because a multidominance approach is pursued, see e.g. Starke 2001, Abels 2004, 2012, Frampton 2004).

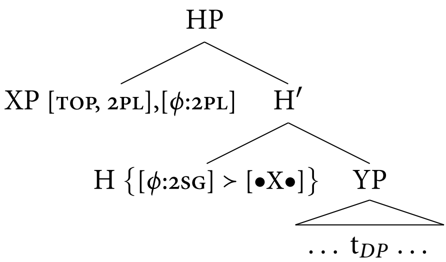
Given these assumptions, I will now show how IM and Agree can interact opaquely. An abstract counter-bleeding configuration is shown in (14) and (15). Suppose there is a head H which triggers IM and ϕ-Agree. For concreteness, it attracts an DP with a topic feature [top]. H also wants to Agree in ϕ-features with the topic DP, which bears the features 2nd person plural. If H first triggers IM as in (14), it moves the DP with the topic feature to its specifier ('[A] > [B]' on H indicates that the operation-inducing feature [A] is discharged before the feature [B]). Afterwards, the DP is no longer in the c-command domain of H. As a consequence, the ϕ-probe on H, which seeks for a goal in the next step, cannot target the DP with the topic feature anymore. It could have done so before the DP moved; hence, topic movement applies too early and thereby destroys the context for the application of Agree between H and $DP_{[TOP]}$. A default value (3sg) is thus inserted on the probe. This bleeding effect can be read off of the surface representation of HP: The DP with the topic feature is not in the c-command domain of H in (14) and hence it is clear why there is no ϕ-Agree. This effect will occur with final movement steps like e.g. topic movement or wh-movement to a head with a [•TOPIC•]- and a [•WH•]-feature, respectively.

(14) *Bleeding*:



However, in the same language that exhibits such bleeding effects, there are also configurations with the topic DP in SpecH but with full ϕ-agreement on H, cf. (15). Given (14), we would expect bleeding of ϕ-Agree here as well, but it does not occur. This instance of counter-bleeding can be derived if the probe and the structure-building feature on H are discharged in the reverse order, i.e., if the DP with the topic feature moves *after* Agree has applied (cf. the order [$*\phi*$] ≻ [•X•] on H in (15)). The consequences are the following: First, the probe seeks for a goal. At this point of the derivation, the topic DP is still in the c-command domain of H and therefore H is valued with the ϕ-features of that DP. Afterwards, H triggers movement of the topic DP.

(15) *Counter-Bleeding*:



On the surface, it is unclear why Agree between H and the topic DP has applied given that the DP is not in the right structural position to be a goal for H, i.e. it is not in H's c-command domain. The explanation is that at the point of probing, the DP has indeed been in the relevant domain, but this is disguised by the subsequent movement of the DP. This counter-bleeding pattern will arise with intermediate movement steps in a movement chain, viz., with movement steps that are triggered by edge features.

The conclusion is that in languages with both types of interactions in the HP, IM sometimes applies before and sometimes after Agree, depending on the nature of the landing site of the moved XP. The landing sites can be locally distinguished on H on the basis of their trigger: Intermediate movement steps are triggered by edge features on H; they are discharged *after* probe features; final movement steps are triggered by structure-buiding features on H that are not categorially underspecified; these are discharged *before* probe features. This analysis presupposes a strictly derivational model of grammar.

4. Case Studies

In this section I present three phenomena that instantiate the abstract pattern of an opaque interaction of Merge and Agree introduced in sections 1 and 3. The counter-bleeding pattern arises with the Anti-agreement effect for ϕ-Agree, and with the possessor case split in Hungarian for case Agree. Defective Intervention in an Icelandic dialect exhibits a counter-feeding interaction with respect to ϕ-Agree.

4.1. (Counter-)Bleeding: The Anti-Agreement Effect

4.1.1. Data and Rule Interactions

In a number of unrelated languages such as Berber, Breton, Welsh, Kinande, Kikuyu, Palaun and Turkish, ϕ-Agree between the verb and the subject reduces to default agreement (3sg) if the subject is Ā-moved (questioned, relativized, focussed) to the SpecC position of the minimal clause. This phenomenon is known as the Anti-agreement effect (AAE, cf. Ouhalla 1993, Richards 1997, Phillips 2001 for an overview). It is illustrated with wh-extraction data from Berber in (16). If the subject is not Ā-moved, the verb agrees with it in person, number and gender, cf. (16-a). If, however, the subject is questioned (Ā-moved to the minimal SpecC), as in (16-b), the verb shows default agreement (3sg masculine, glossed as 'participle').[3] Full agreement is ungrammatical in this context, cf. (16-c). Interestingly, in some of

[3]The verb form in (16-b) is glossed as 'participle', but Ouhalla (1993) argues that this form contains the default 3rd person masculine form of agreement. In most other languages with

the languages that exhibit the AAE when the subject is extracted to SpecC of the minimal clause, long-distance Ā-movement of the subject to SpecC of a higher clause does not result in default agreement in the embedded clause from which the subject is extracted; instead, there is full ϕ-agreement on the embedded verb (cf. (16-d)), as if the subject was not extracted at all.

(16) *Anti-agreement in Berber, wh-movement (Ouhalla 1993: 479f.)*:

a. zri-n imhdarn Mohand
saw-3PL students Mohand
'The students saw Mohand.' *no Ā-IM, full agr.*

b. man tamghart ay yzrin Mohand
which woman COMP see.PART Mohand
'Which woman saw Mohand?' *local Ā-IM, default agr.*

c. *man tamghart ay t-zra Mohand
which woman COMP 3SG.FEM-saw Mohand
'Which woman saw Mohand?' *local Ā-IM, full agr.*

d. man tamghart ay nna-n qa t-zra Mohand?
which woman COMP said-3PL that 3SG.FEM-saw Mohand
'Which woman did they say saw Mohand?' *long Ā-IM, full agr.*

In what follows, I will use the term "short" Ā-movement for Ā-movement to the SpecC position of the minimal clause and "long" Ā-movement for Ā-movement to the SpecC position of a higher clause.

In terms of rule interaction, the AAE can be described as follows: Short Ā-movement of the subject DP *bleeds* ϕ-Agree whereas long Ā-movement has the opposite effect. If a subject that is to undergo long-distance extraction uses SpecC of the minimal clause as an intermediate landing site, we have an instance of *counter-bleeding* when looking at the minimal CP: Under both short and long Ā-movement, there is a DP in the local SpecC position, but sometimes IM of the subject bleeds ϕ-Agree in that position, and sometimes it does not bleed ϕ-Agree in the same position. Put differently, ϕ-Agree did apply in the embedded clause under long extraction, although its context does not seem to be met, given that a DP in the minimal SpecC position can bleed agreement (under short extraction). Hence, on the surface, there is opacity. Importantly, opacity only arises if the subject DP that is to undergo long-

the AAE, the verb form under short Ā-extraction is completely identical to 3rd person singular agreement.

extraction makes a stop-over in SpecC of the embedded clause. For syntactic, morphological and semantic evidence for the successive-cyclic nature of long Ā-movement through the embedded SpecC see McCloskey (1979), Cole (1982), Clements et al. (1983), Torrego (1984), Barss (1986), Lebeaux (1990), Chung (1994), Fox (2000), Lahne (2008*b*) among many others. I adopt this assumption and present another argument in favor of successive-cyclic movement through intervening SpecC positions in AAE languages in section 4.1.3.

4.1.2. Analysis

First of all, I assume that subject-verb agreement in AAE languages is mediated by C and not by T (see also Ouali 2008 and Henderson (2009) for this assumption about some AAE languages).[4] The reason for this assumption is that short and long Ā-movement can be distinguished by the nature of their trigger on the minimal C head: With respect to short Ā-movement of DP, SpecC of the minimal clause is the final landing site for DP; this movement step is triggered by a wh-feature [•wh•] (or [•rel•], which triggers operator movement in a relative clause, etc.). For long Ā-movement, this position is only an intermediate landing site and movement to it is triggered by an edge feature [•X•]. The opacity effect is derived if final movement steps to SpecC apply before Agree and intermediate movement steps to the minimal SpecC apply *after* Agree:

(17) *Order of features on C in Berber type languages*:
C { [•wh•] ≻ [$*\phi*$] ≻ [•X•] }

The consequences of this ordering are as follows: If there is no subject extraction, as shown in (18), C only bears a ϕ-probe and no structure-building features that trigger IM. When the probe searches for the closest goal in its

[4] Since I assume that the ϕ-probe in AAE languages in on C, but the inflection shows up on the verb, there must be a connection between C and V: Either the verb picks up the inflection in C by V-to-C movement (cf. Sproat 1985, Jouitteau 2005 on Celtic) or the inflection is lowered to the T-V-complex in the morphological component. Lowering of the inflection is reminiscent of Affix-Hopping (Chomsky 1957). Furthermore, it seems plausible given the concept of Feature Inheritance, i.e., ϕ-feature transfer from C to T (cf. Chomsky 2004, Richards 2007); the only difference is that I assume that this transfer may apply late, in the morphological component (cf. morphological merger in Embick and Noyer 2001). See also Ouali (2008) on variation in the timing and mode of Feature Inheritance.

c-command domain, it finds the subject DP and is valued with the ϕ-features of this DP.

(18) *No Ā-extraction of the subject DP, ϕ-probe discharged*:
$[_{\text{CP}}\ \text{C}_{\{[*\phi*]\}}\ [_{\text{TP}}\ \text{DP}\ [_{\text{T}'}\ \ldots\ ]]]$
└ ✓Agree ┘

If the subject is questioned in the minimal CP (short Ā-movement), C bears a ϕ-probe and the IM trigger [•WH•]. The structure-building feature is discharged first; it attracts the wh-subject to SpecC, cf. (19-a). Afterwards, the probe initiates Agree but since the closest goal, the subject DP, is no longer in C's c-command domain due to prior movement, Agree fails and the probe is valued by default, cf. (19-b). Hence, final movement of the subject DP to SpecC applies too early and thereby bleeds ϕ-Agree.

(19) *Short Ā-extraction of the subject DP*:

a. Step 1: Movement of the subject to SpecC, wh-feature discharged:
$[_{\text{CP}}\ \text{DP}_{wh}\ \text{C}_{[\bullet\text{WH}\bullet]\ \succ\ [*\phi*]\}}\ [_{\text{TP}}\ \text{t}_{DP}\ [_{\text{T}'}\ \ldots\ ]]]$
└── Move ──┘

b. Step 2: ϕ-Agree initiated by C, default valuation of the probe:
$[_{\text{CP}}\ \text{DP}\ \text{C}_{\{\not[\bullet\text{WH}\bullet]\ \succ\ [*\phi*]\}}\ [_{\text{TP}}\ \text{t}_{DP}\ [_{\text{T}'}\ \ldots\ ]]]$
└ *Agree ┘

In contrast to that, long Ā-extraction that uses the minimal SpecC as an intermediate landing site is triggered by an edge feature. This movement applies late enough and therefore does not bleed Agree: First, the probe on the C head of the embedded sentence initiates Agree and finds the subject DP as its closest goal, cf. (20-a). Only afterwards does the subject move to SpecC, triggered by the edge feature on C, cf. (20-b). On the surface, the subject DP is no longer in the c-command domain of the embedded C, which gives rise to opacity; however, the subject was in the relevant domain at the point of ϕ-Agree.

(20) *Long Ā-extraction of the subject DP*:

a. Step 1: ϕ-Agree initiated by C, ϕ-probe discharged

$[_{CP}\ C_{\{[*\phi*] \succ [\bullet X \bullet]\}}\ [_{TP}\ DP_{wh}\ [_{T'} \ldots]]]$

└── ✓Agree ──┘

b. Step 2: Movement of the subject DP to the local SpecC, edge feature discharged

$[_{CP}\ DP_{wh}\ [_{C'}\ C_{\{\not[\not*\not\phi\not*\not] \succ [\bullet X \bullet]\}}\ [_{TP}\ t_{DP}\ [_{T'} \ldots]]]]$

└──── Move ────┘

Before closing this section, I will address some remaining issues that arise in the derivations presented above. The first question concerns short Ā-extraction in (19). Since short extraction applies early, the ϕ-probe on C cannot target the extracted subject DP anymore. But why can't it just target the closest target that is still in its c-command domain, i.e., the object DP if the verb is transitive? I propose that this is impossible due to the Activity Condition: The object is assigned case by v very early in the derivation. It thereby becomes inactive, i.e., invisible for a ϕ-probe. C thus cannot target the object.[5]

A further issue that arises is that of case assignment to the subject DP: If ϕ-Agree and case assignment go hand in hand, where does a subject that undergoes short Ā-movement to the minimal SpecC get its case value from? Since Agree between C and the subject DP is bled in this configuration C can also not assign a case to the subject. I propose that default valuation applies for case just as for ϕ-features. The case feature of the subject is valued with the default nominative case. See section 4.3 for another instance of default case valuation under short extraction.

[5]In section 4.2.3 I will argue that SpecT is an intermediate landing site for movement to SpecC. If this holds in general, as I will assume, a related problem arises: The AAE does not occur with long Ā-movement of the object. If, however, a wh-moved object goes through SpecT, there is a stage in the derivation at which the wh-object is in SpecT of the embedded clause and above the subject DP. Why can the ϕ-probe on C not Agree with the closer object DP in this configuration (since the object undergoes long extraction and checks the edge feature on the embedded C, it is still in C's c-command domain when the ϕ-probes searches for a goal)? As proposed above, this is impossible because the object is inactive for ϕ-Agree at that point of the derivation.

4.1.3. *More Evidence from Variation for SpecC as an Intermediate Landing Site*

The AAE data are only opaque if long Ā-movement makes a stop-over in the embedded SpecC position. However, there is an alternative analysis of the counter-bleeding effect in AAE languages that does not rely on this intermediate movement step: If short movement of a DP to SpecC causes bleeding but there is no bleeding under long Ā-extraction, one could propose that wh-movement to the matrix clause does not make a stop-over in the SpecC position of the embedded clause but moves directly to the matrix SpecC. Under this analysis, it follows why we find full subject-verb-agreement in the embedded clause: When the ϕ-probe on the embedded C searches for a goal, the wh-subject is still in its base position in the embedded clause and thus in the embedded C's c-command domain; it can only move when matrix C is merged with the IM trigger [•WH•]. But due to incremental bottom-up structure-building, matrix C is merged later than the embedded C head. This account does not require ordering of two different types of IM.

However, the assumption that long Ā-movement does not go through the embedded SpecC position causes problems for languages in which *both* long and short Ā-movement bleed ϕ-Agree. This is the case for example in some Italian dialects (Fiorentino and Trentino, cf. Brandi and Cordin 1998, Campos 1997) and in Ibibio (cf. Baker 2008). (21) illustrates that in Fiorentino full subject agreement under short extraction is ungrammatical; instead, default agreement must be used, compare (21-a) and (21-b). Default agreement is also necessary under long subject extraction, compare (21-c) and (21-d).

(21) *The AAE in Fiorentino (taken from Ouhalla 1993, Campos 1997)*:

a. Quante ragazze gli ha parlato con te?
how.many girls CL.3SG have.3SG spoken to you
'How many girls (it) has spoken to you?'

b. *Quante ragazze le hanno parlato con te?
how.many girls CL.3PL have.3PL spoken to you
'How many girls have spoken to you?'

c. Quante ragazze tu credi che gli ha telefonato?
how.many girls you think that CL.3SG have.3SG phoned
'How many girls do you think have phoned?'

d. *Quante ragazze tu credi che le hanno telefonato?
how.many girls you think that CL.3PL have.3PL phoned

In contrast to Berber, long Ā-movement of the subject also bleeds Agree in the embedded clause. But under the assumption that long extraction does not go through the embedded SpecC position, this is mysterious: The subject DP that is to undergo long extraction is still in the c-command domain of C when the latter probes and should thus value the probe. One solution to this problem could be to say that languages differ in the size of their locality domains: In AAE languages of the Berber type, long Ā-movement does not go through SpecC, which explains the absence of bleeding with long extraction; in AAE languages of the Trentino type, however, long Ā-movement does go through SpecC and DPs in the embedded SpecC cause bleeding (due to early movement prior to ϕ-Agree), regardless of whether SpecC is their final landing site or not.

In my view it is clearly undesirable to assume that languages differ in the size of their locality domains. In the languagues in which long Ā-movement does not apply successive-cyclically, there must be a non-local dependency between the matrix C and a DP in the embedded clause. This goes against the trend in contemporary minimalist syntax, where non-local dependencies are decomposed into a series of local dependencies (cf. Alexiadou et al. 2012). Under the analysis of the opacity effect presented in 4.1.2, there is another way to account for variation between AAE languages that I will adopt: Assume that languages do not have different locality domains for movement. In order to model bleeding with long extraction in the Trentino type languages, the embedded SpecC must be an intermediate landing site. So it must be in languages of the Berber type, with the consequence that the AAE data in these languages are indeed opaque. Cross-linguistic variation is accounted for by reordering of operation-inducing features on C: In the Trentino type languages, *both* edge-feature-driven IM and non-edge feature-driven IM apply *before* Agree. Hence, the wh-subject DP is moved out of the c-command domain of C before probing happens and there is thus always bleeding, regardless of the IM type.

(22) *Order of features on C in Trentino type languages*:
C { [•WH•], [•X•] ≻ [∗ϕ∗] }

In Berber type languages, on the other hand, the two IM triggers apply at different points in the derivation relative to Agree, giving rise to opacity; compare the order of features in (17) and (22).

Previous analyses of the AAE pattern include Ā-binding approaches (Brandi and Cordin 1998, Ouhalla 1993) and anti-locality approaches (Cheng 2006, Schneider-Zioga 2007). The basic idea of the former type of approach is that subject extraction leaves behind a *pro* in the subject position that must not be bound by its antecedent from an Ā-position (Principle B applied to Ā-binding). It is, however, bound under short subject extraction by the subject in SpecC. In order to avoid the violation of Principle B, *pro* must not be licensed in the first place. It is licensed by rich subject-verb-agreement; dropping the agreement, resulting in AAE, makes *pro* unavailable and thus avoids the principle B violation.

In the anti-locality approaches, it is argued that short extraction of the subject is too local in the sense of Grohmann (2003) (at least in some AAE languages). Hence, a repair mechanism must apply: The lower copy of the subject is spelled out, but not as a DP or a pronoun; rather, it is the anti-agreement morphology (default agreement) that is the spell-out of that copy.

Apart from a number of conceptual problems with these approaches that I will not address here, both of the approaches have problems to account for cross–linguistic variation. One approach needs to postulate that languages differ in the size of their locality domains, and the other must assume that they differ in the properties of empty elements such as traces. In the Ā-binding approach presented in Ouhalla (1993), long Ā-extraction passes through SpecC of the embedded CP. Since languages like Berber do not show the AAE with long extraction, Ouhalla assumes that the trace of the subject in the embedded SpecC position cannot act as an Ā-binder. To account for the occurrence of the AAE with long extraction in the Trentio-type languages, he proposes that the trace can be an Ā-binder in these languages. Hence, languages vary in the ability of traces to be Ā-binders. Apart from the question whether movement leaves behind traces, copies or maybe no element at all, it seems questionable that an empty element left by the same movement operation in the same position varies in this way.

In the anti-locality approaches, it is simply assumed that long extraction does not go through the embedded SpecC position but moves directly to the matrix SpecC (in languages of the Berber type). Since this movement is not too local, there is no repair, i.e., no anti-agreement morphology on the embedded verb. The Trentino pattern with anti-agreement showing up even under long extraction is not addressed in these approaches. One would probably have to assume that long movement does make a stop-over in the embedded

SpecC position. Hence, languages have to differ in their locality domains. As pointed out above, I take this to be undesirable. The reordering approach presented in this paper does not have to assume that languages differ in this respect, and it also does not need to postulate that elements left by movement (if there are such elements at all) have different binding properties. I take this to be a desirable outcome.

4.2. (Counter-)Feeding: Intervention Effects in Icelandic B

4.2.1. Data and Rule Interactions

In many languages, a dative XP blocks Agree relations between a probe on a head H that c-commands the dative XP and a lower DP that is c-commanded by H and the dative XP. This is remarkable because the dative XP itself can also not value the probe on H. This effect is known as defective intervention (Chomsky 2001). In Icelandic, dative experiencers (Exp) are defective interveners for Agree between T and a lower subject DP. As shown in (23-a), the raising verb can only bear default 3rd singular agreement because Agree with the subject in the embedded infinitive is blocked by Exp. Sigurðsson and Holmberg (2008) describe three Icelandic dialects that pattern differently with respect to the presence vs. absence of the intervention effect in the context of experiencer movement. Opaque interaction of Agree and IM is found in Icelandic B for number agreement as shown in (23-b) and (23-c) (see also Holmberg and Hróarsdóttir 2003 for a description of that dialect):[6]

(23) *Raising constructions in Icelandic B (Holmberg and Hróarsdóttir 2003)*:

a. það virðist/*virðast **einhverjum manni** [hestarnir
there seem.3SG/seem.3PL some man.DAT the-horses.NOM
vera seinir]
be slow

[6] As Sigurðsson and Holmberg (2008) show, person agreement patterns differently from number agreement. Agreement in 1st and 2nd person is blocked in all three dialects, regardless of the position of the dative experiencer. In what follows, I confine myself to number agreement. The facts on person agreement can be integrated into the present analysis if person and number are separate probes, as Sigurðsson and Holmberg (2008) have argued, and if the person probe, searching for local person arguments, probes *before* any movement operation has taken place; i.e., the person probe is the feature on T which is discharged before all other operation-inducing features.

‘It seems to some man that the horses are slow.’

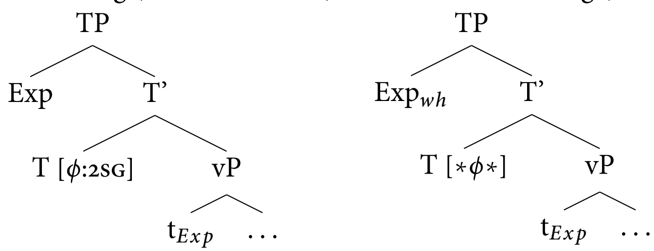
b. **Mér** virðast t_{NP} [hestarnir vera seinir]
me.DAT seem.3PL the-horses.NOM be slow
‘It seems to me that the horses are slow.’

c. **Hvaða manni** veist þú að virðist/*viðast t_{wh}
which man.DAT know you that seem.3SG/seem.3PL
[hestarnir vera seinir]
the-horses be slow
‘To which man do you know that the horses seem to be slow?’

If the experiencer moves above the ϕ-probe on T in the matrix clause in order to check the EPP feature on T, this movement step feeds Agree between T and the subject DP in the embedded clause, cf. (23-b).[7] This is not surprising since, on the surface, the experiencer does not intervene anymore between probe and goal. What is striking, however, is that if the experiencer is moved to SpecC (questioned, relativized, topicalized), it does not feed Agree, although on the surface it also does not intervene anymore, cf. (23-c). This is a case of counter-feeding: Agree between T and the subject DP cannot apply although its context is given – the intervener is moved out of the way, just as in (23-b). Hence, a wh-moved experiencer behaves as if it is not moved at all with

Opacity arises at the stage of the derivation where TP is complete, cf. (24). Assume that wh-movement to SpecC makes a stop-over in SpecT (see section 4.2.3 for justification). At this stage in both (23-b) and (23-c), the experiencer is in SpecT. But in (23-b) it feeds Agree in this position whereas in (23-c) it has the opposite effect in the same position.

(24) a. Feeding (EPP-movement): b. Counter-feeding (wh-movement):



[7]Due to the V2-property of Icelandic, the finite verb moves to C and the dative subject in (23-b) undergoes further movement to SpecC in order to fill the sole preverbal position. What is important for the present discussion is that the subject has also checked the EPP in SpecT (see Holmberg and Hróarsdóttir 2003, Sigurðsson and Holmberg 2008 for evidence); subject movement to SpecT is thus a final movement step.

4.2.2. Analysis

The crucial difference between the feeding and the counter-feeding example again lies in the landing site of the experiencer: In (23-b), SpecT is the final landing site in the movement chain triggered by the EPP on T.[8] In (23-c), however, SpecT is only an intermediate landing site on the way to SpecC, where the feature [•WH•] is checked (for discussion of how the EPP is checked in this configuration see below). Intermediate movement steps are triggered by an edge feature on T and apply *after* Agree; final movement steps to SpecT are triggered by the EPP feature, represented as [•D•] in what follows, and apply *before* Agree. Just as in the AAE languages of the Berber type, the two movement types apply at different points in the derivation with respect to Agree.

(25) *Order of features on T (Icelandic B)*:
T { [•D•] ≻ [$*\phi*$] ≻ [•X•] }

The consequences of this ordering are the following: If the experiencer is not moved at all and the EPP is checked by another element such as the expletive in (23-a), the experiencer is in the c-command domain of T and thus intervenes for ϕ-Agree between the probe on T and the lower subject DP, cf. (26).

(26) *No movement of the experiencer: default agreement*:
[$_{\text{TP}}$ T$_{\{[*\phi*]\}}$ [$_{\text{vP}}$ v [$_{\text{VP}}$ Exp [$_{\text{V}'}$ V [$_{\text{TP}}$ DP [$_{\text{T}'}$ …]]]]]]
*Agree (T → DP)

If the experiencer undergoes EPP-movement, this movement step applies before T initiates Agree, cf. (27-a). Afterwards, the ϕ-probe on T can target the subject DP since the experiencer is no longer in the c-command domain of T and thus does not intervene anymore, cf. (27-b).

(27) *EPP-movement of the experiencer, full agreement*:

a. Step 1: movement of the experiencer to SpecT, EPP discharged
Move (Exp → SpecT)
[$_{\text{TP}}$ □ [$_{\text{T'}}$ T$_{\{[\bullet D\bullet] \succ [*\phi*]\}}$ [$_{\text{vP}}$ Exp [$_{\text{v'}}$ v [$_{\text{VP}}$ t$_{Exp}$ [$_{\text{V}'}$ V [$_{\text{TP}}$ DP [$_{\text{T}'}$ …]]]]]]]]

[8] SpecT is not the ultimate landing site for the subject, which undergoes further movement to SpecC. Recall that "final" means final in a given chain. Movement to SpecT is the final step in the movement chain created to check the EPP in SpecT. See section 2.3 and footnote 7.

b. Step 2: Agree between T and the lower DP, ϕ-probe discharged:
[$_{TP}$ Exp [$_{T'}$ T$_{\{}$~~[•D•]~~ ≻ [$*\phi*$]$_{\}}$ [$_{vP}$ t$_{Exp}$ [$_{v'}$ v [$_{VP}$ t$_{Exp}$ [$_{V'}$ V [$_{TP}$ DP [$_{T'}$ …]]]]]]]]
✓Agree

If the experiencer is questioned and must check the [•WH•]-feature on C, it uses SpecT only as an intermediate landing site. This movement step is triggered by an edge feature and applies after Agree. First, the ϕ-probe on T searches for a goal but Agree is blocked because the wh-experiencer still intervenes, cf. (28-a). It is moved out of the way only afterwards, cf. (28-b). Movement of the experiencer comes too late to feed Agree.

(28) *Wh-movement of the experiencer, default agreement*:

a. Step 1: Agree initiated by T fails, Exp$_{wh}$ still intervenes, default agreement
*Agree
[$_{T'}$ T$_{\{}$ [$*\phi*$] ≻ [•X•]$_{\}}$ [$_{vP}$ Exp$_{wh}$ [$_{v'}$ v [$_{VP}$ t$_{Exp}$ [$_{V'}$ V [$_{TP}$ DP [$_{T'}$ …]]]]]]]

b. Step 2: Intermediate movement step to SpecT, edge feature discharged:
Move
[$_{TP}$ Exp$_{wh}$ [$_{T'}$ T$_{\{}$ ~~[$*\phi*$]~~ ≻ [•X•]$_{\}}$ [$_{vP}$ t$_{Exp_{wh}}$ [$_{v'}$ v [$_{VP}$ t$_{Exp}$ [$_{V'}$ V [$_{TP}$ DP [$_{T'}$ …]]]]]]]]

With respect to the counter-feeding derivation in (28) a question arises: To obtain counter-feeding it is important that the wh-experiencer checks the edge feature on T and not the EPP feature. If it could check the EPP, movement to SpecT would apply early and would lead to feeding with wh-movement. However, since the EPP is always present on T (in contrast to the [•WH•]-feature on the embedded C head in the AAE derivation with long Ā-extraction), it is not clear why a wh-phrase cannot check it. I do not have an answer why the wh-phrase does not discharge the EPP, but it is clear for Icelandic that it really doesn't: As Holmberg and Hróarsdóttir (2003) show, the EPP can be checked by a different phrase in exactly the configuration where the experiencer undergoes wh-movement (or relativization, topicalization) to SpecC. This is an instance of stylistic fronting, an operation that moves the closest overt VP-internal category to SpecT to check the EPP (Holmberg 2000). Stylistic fronting is optional. An instance of this operation in the context of a wh-experiencer is given in (29) where *Òlafur* is fronted (Holmberg and Hróarsdóttir 2003: 1009):

(29) Hverjum hefur Òlafur virst vera gàfaður?
whoDAT has OlafNOM seemed be intelligent
'To whom has Olaf seemed to be intelligent?'

4.2.3. *More Evidence from Variation for SpecT as an Intermediate Landing Site*

On the surface, the interaction of Agree and IM in Icelandic B is opaque because, when looking at the output structure, neither the EPP-moved nor the wh-moved experiencer intervenes between the matrix T and the embedded subject; nevertheless, the wh-experiencer behaves as if it intervened. Derivationally, this could be derived by assuming that the wh-experiencer moves directly to SpecC without a stop-over in SpecT. Under this analysis, it is clear why the wh-experiencer causes an intervention effect: When T probes, the wh-experiencer is still in the c-command domain of T and thus blocks Agree. It is moved to SpecC only after C is merged (this is the solution proposed by Holmberg and Hróarsdóttir 2003). However, assuming that wh-movement does not stop in SpecT causes problems for cross-linguistic variation found with defective intervention: In another Icelandic dialect, Icelandic A, *both* wh-movement and EPP-driven movement of the experiencer feed Agree, i.e., (23-c) is grammatical with 3rd person plural agreement on the verb in this dialect (Sigurðsson and Holmberg 2008). But wh-movement of the experiencer should not cause feeding if wh-movement does not go through SpecT. When T probes for the lower DP, a wh-experiencer still intervenes; it is moved when C is merged with TP at a later stage of the derivation.

A similar pattern exists in Romance languages and Greek, although feeding cannot be read off of the agreement on the verb (McGinnis 1998, Anagnostopoulou 2003): In these languages, an experiencer blocks movement of (instead of ϕ-Agree with) a lower DP to SpecT. If the experiencer is cliticized (adjoined to T) or questioned (moved to SpecC), movement of the lower DP to SpecT becomes possible (feeding). Hence, both types of IM, i.e., movement to SpecC and to the T domain, feed Agree. Assuming that wh-movement does not go through SpecT is problematic for analyzing these languages as well: When T tries to attract the lower DP, a wh-experiencer still intervenes because it is only moved after C is merged, which only happens after the TP is complete. A way out would be to assume that in both Icelandic A, Greek and the Romance languages, direct wh-experiencer movement to SpecC must precede attraction of the subject DP, but the latter step is counter-cyclic if a wh-Exp does not go through SpecT. For Greek and the Romance languages, several solutions to this problem have been proposed in the literature: (i) phase-internal counter-cyclicity (Anagnostopoulou 2003), (ii) intervention is evaluated at the phase-level (McGinnis 2001), (iii) covert movement of the wh-experiencer to a low Ā-position in the c-command domain of T that is invisible to a probe on T because ϕ-Agree is an A-relation and therefore cannot target elements in Ā-positions (Legate 2002). Solution (i) gives up the Strict Cycle Condition within a phase (CP and vP) and thus standard locality assumptions. Solution (ii) suggests that intervention is computed when a phase is complete, i.e., at the CP level. Since a wh-experiencer is in SpecC at the end of derivation, it does not intervene on the surface. However, this means that the Minimal Link Condition is a strongly representational constraint (for the examples under discussion, it is evaluated at the end of the derivation) in a system that is assumed to be strictly derivational. Hence, it is argued that the data are not compatible with a strictly derivational model of grammar (this has also been pointed out in Řezač 2004). Finally, solution (iii) resorts to a covert movement step to a low Ā-position that is hardly motivated independently. In order to avoid all these problematic assumptions one could propose that languages differ in the size of their locality domains: In Icelandic B-type languages, wh-movement does not make a stop-over in SpecT, which accounts for the absence of feeding with wh-experiencers. In languages of the IcelandicA-/Romance/Greek type, however, wh-movement must proceed in smaller steps and does go through SpecT, explaining the feeding effect. Just as with the variation found with the AAE, the question is whether it is desirable to assume that languages differ in such a fundamental property like the locality domains for movement.

I will pursue another solution that (i) neither needs to weaken the Strict Cycle Condition nor (ii) needs to invoke strongly representational constraints (evaluated at the end of the derivation), nor (iii) postulates different locality domains in different languages. I assume that wh-movement always goes through SpecT, in all the languages discussed above (for the same assumption cf. Chomsky 2005, Gallego 2006, see also Müller 2004 and references cited there). This is needed to account for feeding in Romance-type languages and since languages do not vary in this property, it must hold for languages of the Icelandic B-type as well (and the data from this dialect are thus indeed opaque). As with the variation in AAE languages, I propose that cross-linguistic variation is accounted for by reordering of operation-inducing features on a head. In contrast to Icelandic B, the EPP and the edge feature are ordered before the probe feature in T in Icelandic A and Romance/Greek, cf. (30).

(30) Order of features on T in Icelandic A:
T { [•D•], [•X•] ≻ [$*\phi*$] }

As a consequence, the experiencer, whether it is a wh-experiencer or not, moves early to SpecT and does not intervene anymore at the point when T probes. In Icelandic B, the two IM triggers align differently with respect to the probe feature, which gives rise to opacity. In this way, the transparent (Romance, Icelandic A) and opaque (Icelandic B) interactions of Merge and Agree in the two types of languages can be accounted for in a uniform way. The only difference between them is the order of operation-inducing features on a head that triggers more than one operation – a purely lexical property. Languages do not differ in the constraints that determine the size of locality domains (like the PIC). The approach has the additional advantages that standard locality constraints are obeyed to and that the strongly representational version of the MLC as proposed in Anagnostopoulou (2003) can be avoided in a strictly derivational system.[9]

[9]Fabian Heck has pointed out to me that if the intervention of the dative experiencer in the present analysis is computed on the basis of an MLC-like constraint, the MLC would still have a representational residue because it evaluates the distance of two potential goal XPs relative to the T head, i.e., the TP must be scanned to see which goal is closer to T; however, this version of the MLC is arguably of a weaker type because it is evaluated before the derivation is complete, in contrast to the strongly representational version discussed in the main text. Note that the portion of the derivation that needs to be accessible to evaluate intervention can be further reduced if a probe simply targets the first available goal in its c-command domain and if the dative experiencer actually counts as a potential goal which can, however, not value the probe (as has been suggested in the literature on defective intervention, see Preminger 2011 for an overview of relevant accounts). As a consequence, the probe stops searching as soon as it finds the experiencer in its c-command domain and does not access material that is lower down in the structure.

To summarize, the Icelandic B data are not only opaque when looking at the surface position of the experiencer relative to the T head; they are also opaque when looking at the TP (before the derivation is complete): The experiencer is in SpecT, regardless of whether it checks the EPP or an edge feature in this position, but movement of the experiencer to SpecT has different consequences for Agree. Hence, Agree and Internal Merge interact opaquely.

4.3. (Counter-)Bleeding: Possessor Case and Agreement in Hungarian

4.3.1. Data and Rule Interactions

In this subsection I will present another instance of counter-bleeding. In contrast to the two previous instances of opaque interactions of Merge and Agree, the Agree relation will result in case valuation instead of ϕ-agreement.

The phenomenon that shows opacity in this respect is the case marking of the possessor in Hungarian. In Hungarian, the possessor precedes the possessum and the possessum agrees in ϕ-features with the possessor. The possessor exhibits a case split: It can bear either nominative or dative, cf. (31). As argued in Szabolcsi (1994), the two possessors occupy different structural positions: the dative possessor is in an operator position within the DP, which is taken to be SpecD. It is moved to that position from a lower position that is associated with nominative case. Evidence for this analysis comes from two facts: (i) the dative possessor precedes the D element *a(z)* whereas the nominative possessor follows this element (cf. (31-a) vs. (31-b)); (ii) only the dative possessor can be extracted out of the DP, cf. (32). As long as the possessor is not to undergo movement out of the DP, the choice between a nominative and a dative possessor is optional. Finally, note that both the dative and the nominative possessors precede a special class of determiners like 'each', cf. (31-c).

(31) *Possessors in Hungarian*:

a. (a) Mari kalap-ja
the Mari.NOM hat-POSS.3SG
'Mari's hat'

b. Mari-nak a kalap-ja
Mari.DAT the hat-POSS.3SG
'Mari's hat'

c. a te valamenyi tiik-od
the you.NOM each secret-POSS.2SG
'your every secret'

(32) *Extraction data*:

a. Mari-nak nem ismert-em [t′ t növér-é-t]
   Mari.DAT not knew-1SG sister-POSS.3SG-ACC
   'I never knew any sister of Mari'

b. *Mari nem ismert-em [t′ t növér-é-t]
   Mari.NOM not knew-1SG sister-POSS.3SG-ACC
   'I never knew any sister of Mari.'

With this empirical background, I make the following assumptions about the structure of the DP in Hungarian, illustrated in (33): The possessor (Poss) is merged as a sister of of N (the possessum) for reasons of theta-role assignment (cf. Delsing 1998, de Vries 2006, Georgi and Salzmann 2011). Determiner elements like 'each' are merged on top of NP, i.e., NP is the sister of Det. In order to derive the surface order Poss > Det, the possessor must move to the specifier of a functional projection above Det. I take this to be the projection of the functional head n. n has a trigger for this movement, represented as [•D_{Poss}•] in what follows. In addition, n assigns dative case to the possessor and thus bears a case probe [∗dat∗]. The nominative is not assigned to a possessor; rather, it is the default value instantiated on a possessor if it does not Agree in case with n. Finally, D has a feature that triggers movement of the possessor to SpecD, but only if it bears dative case: [•D_{dat}•] (recall that only dative possessors occupy SpecD, nominative possessors stay in the lower Specn position, preceding Det but following D elements).

(33) *Structure of the Hungarian DP*:

[$_{DP}$ □ [$_{D'}$ D [$_{nP}$ Poss [$_{n'}$ [$_{NP}$ Det [$_{N'}$ ⟨Poss⟩ N]] n]]]]
   ... Move ... ——— Move ———

The following interactions between Agree and Merge arise in the DP: Movement of Poss to Specn is obligatory since all possessors precede Det. Movement to this position may bleed dative case assignment by n, resulting in default nominative case on the possessor. Opacity arises when looking at the nP: Assuming that any Poss must go through Specn, both dative and nominative possessors occupy the same position, but with different consequences for case Agree: In one case dative assignment is bled by movement to Specn; in the other case, however, dative assignment does take place although Poss is in the same position Specn, an instance of counter-bleeding.

4.3.2. *Analysis*

Again, the bleeding and the counter-bleeding case can be distinguished by their trigger: If the possessor ends up with nominative (dative assignment is bled), it stays in Specn; Specn is thus its final landing site. If, however, the possessor moves on to SpecD (dative possessors), Specn is only an intermediate landing site for the possessor. Intermediate movement steps are triggered by an edge feature on n, whereas final movement steps are triggered by [•D_{Poss}•] on n. The opacity effect can be derived if the final movement step applies *before* Agree and the intermediate movement step applies *after* Agree. n-Poss-Agree in ϕ-features is not influenced by the case of the possessor, there is always full agreement in ϕ-features. Since ϕ-agreement is never bled by the movement of the possessor, this suggests that it applies before any movement operation has applied. The ϕ-probe on n thus always finds the possessor as a goal in its c-command domain.[10] The order of features on n is given in (34):

(34) *Order of features on n*:
n { [$*\phi*$] ≻ [•D_{Poss}•] ≻ [∗c:dat∗] ≻ [•X•] }

This order has the following consequences (for the sake of clarity, ϕ-Agree is not indicated, it applies first and always results in valuation of the ϕ-probe): If the possessor ends up with nominative in Specn, its movement is driven by the [•D_{Poss}•] on n. This operation applies early, cf. (35-a). Afterwards, n triggers dative assignment, but since the proper goal, the possessor, is no longer in n's c-command domain, it cannot be assigned dative case, cf. (35-b). The result is default nominative valuation.[11]Finally, D is merged with nP. Since

[10]Note that this is also evidence for the separation of case- and ϕ-Agree in Hungarian since the two apply at different points in the derivation. See also Marantz (1991), Bobaljik (2008), Baker and Vinokurova (2010), Keine (2010), Preminger (2011) among others for this conclusion.

[11]One might wonder why n cannot assign dative case to the possessum which is in its c-command domain. It does not bear a case value at that point of the derivation (and is thus not inactive) since its value is determined by the external head that selects the whole DP. Note that this is a very general question: Often, the possessor case is assumed to be assigned by a head that has both the possessor and the possessum in its c-command domain. It is unclear why it targets the possessor and not the possessum, a problem that a number of approaches face. See Georgi and Salzmann (2011) for extensive discussion and a number of references to papers that encounter the same question with respect to the possessor doubling construction in German. The following solutions come to mind: (i) Case assignment is tied to XPs and not to heads; (ii) dative assignment is tied to a specific property of the possessor DP, e.g. its thematic role; (iii) case stacking: n assigns dative to N, but in addition N is assigned the case from

the possessor does not bear dative, it cannot be attracted to SpecD by the feature [•D_{dat}•] on D. This feature is deleted by default.

(35) *Nominative possessor*:

a. Step 1: EPP-movement of Poss to Specn, EPP discharged:
[$_{nP}$ Poss [$_{n'}$ n $_{[\bullet D_{Poss}\bullet]\succ[*c:dat*]}$ [$_{NP}$ Det [$_{N'}$ t N]]]]
Move

b. Step 2: Agree initiated by n fails, default case on Poss, case-probe discharged:
[$_{nP}$ Poss [$_{n'}$ n ~~[•D•]~~ ≻ [*c:dat*] [$_{NP}$ Det [$_{N'}$ t N]]]]
*Agree

Dative possessors move after n initiated case Agree because this movement is triggered by an edge feature.[12] First, n probes. finds the possessor in its c-command domain and assigns it dative case, cf. (36-a). Afterwards, the possessor undergoes movement to Specn, cf. (36-b). This movement comes too late to cause bleeding since Agree applied earlier. Finally, the dative possessor is moved to SpecD triggered by [•D_{dat}•]. This step is not indicated in (36).

(36) *Dative possessor*:

a. Step 1: Agree, n assigns dative to Poss, case-probe discharged:
[$_{nP}$ n $_{[*c:dat*]\succ[\bullet X\bullet]}$ [$_{NP}$ Det [$_{N'}$ Poss N]]]
✓Agree

b. Step 2: Edge-feature-driven Internal Merge of Poss to Specn:
[$_{nP}$ Poss n ~~[*c:dat*]~~ ≻ [•X•] [$_{NP}$ Det [$_{N'}$ t N]]]
Move

the external head that selects the DP. Only the case assigned by the external head is realized in the postsyntactic morphological component, cf. Moravcsik (1995), Bejár and Massam (1999), Merchant (2006), Assmann, Edygarova et al. (2013).

[12] As with wh-movement of the experiencer in Icelandic where wh-Exp does not check the EPP on T, the question arises why a (future) dative possessor cannot check the feature [•D_{Poss}•] on n. If it did, it would move early and bleed dative case assignment. I do not have an answer to this question. In contrast to Icelandic, no other element seems to check that feature in this configuration.

The abstract pattern of interaction is identical to the one in AAE languages: Intermediate movement steps counter-bleed Agree because they apply after Agree; final movement steps bleed Agree because they apply before Agree. The only difference between the AAE and possessor case in Hungarian is the functional head that triggers IM and Agree, as well as the morphological reflex of the Agree relation. It is head-marking (agreement) in AAE languages and dependent-marking (case) in Hungarian.

5. Cross-Linguistic Variation

In this section I look at the cross-linguistic variation that the present account predicts for the AAE, defective intervention and the encoding of possessors. I will discuss what the attested variation tells us about (i) the extrinsic / intrinsic dichotomy and (ii) the role of independent principles of the grammar that have been argued to determine rule ordering.

5.1. Variation in AAE, Intervention and Possessor Encoding

In section 3 I assumed that the order of operation-inducing features on a head H that triggers more than one operation is in prinicple free, i.e., languages can vary in the order of features on H; it is determined language-specifically. Reordering of the features thus predicts a certain range of variation. If a head triggers Agree (in ϕ-features in (37)) as well as intermediate and final movement steps, represented by [•X•] and [•F•], respectively, we expect four permutations (ignoring the order between the two types of IM triggers when they both apply before or after Agree because in that case, their order cannot be determined by the presence or absence of the morphological reflex caused by the Agree relation):

(37) *Permutations of a probe feature and the two types of IM triggers*:

a. [•F•], [•X•] ≻ [$*\phi*$] (transparent)
b. [$*\phi*$] ≻ [•F•], [•X•] (transparent)
c. [•F•] ≻ [$*\phi*$] ≻ [•X•] (opaque)
d. [•X•] ≻ [$*\phi*$] ≻ [•F•] (opaque)

In (37-a) and (37-b) internal Merge and Agree interact transparently: Both types of IM feed or bleed Agree because they apply before or after Agree. The interaction of Agree and internal Merge in (37-c) and (37-d) is opaque because one type applies before and the other after Agree; as a consequence, one type feeds or bleeds Agree and the other counter-feeds or counter-bleeds Agree. In the table in (40) I cross-classified the empirical phenomena discussed in this paper and the four ordering patterns; the cells are filled with languages that instantiate the the respective patterns.

The data presented in section 4 are all instances of the pattern in (37-c). In sections 4.1.3 and 4.2.3, several examples of the pattern in (37-a) have been discussed: The data from Trentino type languages, where both long and short Ā-movement bleeds ϕ-Agree, and from Icelandic A or Romance-type languages, in which both EPP- and wh-movement of an experiencer cause feeding, served as an argument for (i) reordering of operation-inducing features in the first place, and (ii) for SpecC and SpecT as intermediate landing sites for long Ā-movement, respectively. I take it to be accidental that I did not find an instance of this pattern for possessor case/agreement.

Another instance of this pattern with possessor extraction can be found in other Uralic languages (cf. Nikolaeva 2002) for ϕ-Agree instead of case Agree: In Nentes (Samoyedic branch of the Uralic languages), for example, it is not the case value of the possessor that varies with final or intermediate movement of the possessor but ϕ-Agree between the possessor and the possessum that shows the same sensitivity to the type of possessor extraction: There is overt agreement with so-called internal possessors but not with external possessors:

(38) *Nenets (Nikolaeva 2002)*:

a. tyuku° Wata-h ti
this Wata-GEN reindeer
'this reindeer of Wata'

b. *tyuku° Wata-h te-da
this Wata-GEN reindeer-3SG
'this reindeer of Wata'

c. Wata-GEN tyuku° te-da
Wata-GEN this reindeer-3SG
'this reindeer of Wata'

External possessors pattern with dative possessors in Hungarian in that they precede the D element and can be extracted out of the DP; internal possessors pattern with nominative possessors in Hungarian in that they follow the

D element and cannot be extracted out of the DP. The case of the possessor is constant; it always bears genitive, regardless of the IM type it undergoes. Nenets is thus very similar to Hungarian. The only difference is that in Nenets it is head-marking (agreement on the possessor) that reflects the difference between the two IM types, whereas in Hungarian is is dependent-marking (case on the possessor). The Nenets pattern is derived if the ϕ-probe and the case probe on n in Hungarian are interchanged:

(39) *Order of features on n in Nenets*:
n { [∗c:gen∗] ≻ [•D_{Poss}•] ≻ [∗ϕ∗] ≻ [•X•] }

As a consequence of this order, the possessor will always be assigned genitive because it is still in the base position when the case probe on n searches for a goal. Final movement to Specn applies before ϕ-Agree and thus bleeds it; hence, internal possessors do not show agreement on the possessum. External possessors that use Specn only as an intermediate landing site move after ϕ-Agree initiated by n and thus counter-bleed ϕ-Agree – this is also an instance of the pattern in (37-c).

Another instance can be found in Selkup (Samoyedic branch of the Uralic languages). Selkup exhibits a combination of the Hungarian and the Nenets system because head-marking and dependent-marking are both sensitive to the IM type of the possessor: Internal possessors neither have an overt case marker nor do they trigger agreement; external possessors bear an overt case marker (locative) and trigger agreement. This is derived under the order in (40) where both Agree operations apply in between final and intermediate movement:

(40) *Order of features on n in Selkup*:
n { [•D_{Poss}•] ≻ [∗ϕ∗], [∗c:loc∗] ≻ [•X•] }

Next, consider the pattern in (37-b). There are examples of this pattern for all three phenomena discussed in this paper. With respect to AAE, there are languages where subject-verb-agreement is not bled by any type of IM. In German or English, for example, there is always full subject-verb-agreement under short and long subject Ā-extraction. As for defective intervention, Sigurðsson and Holmberg (2008) describe a third dialect of Icelandic, Icelandic C, which instantiates this abstract pattern: Neither type of experiencer movement feeds Agree between T and the lower subject DP (with respect to number agreement; on person agreement in defective intervention environments see footnote 6). We have already seen this pattern in the discussion of possessor case in Hungarian: Recall that ϕ-Agree between the possessor and the possessum, which I take to be mediated by n, is never bled by any type of IM of the possessor. This suggests that both IM types apply after ϕ-Agree (cf. (34)) with the possessor in the c-command domain of n has taken place.

Interestingly, the pattern in (40-d) does not seem to be attested for any of the three phenomena. This will be further discussed in the next subsection.

Finally, note that the attested variation sheds new light on the timing of the insertion of edge features: Chomsky (2000, 2001) suggests that edge features are inserted once a head is inactive, i.e., once it has discharged all of its operation-inducing features. The edge feature, if present on a head H, is thus the last feature that is discharged on H. In contrast, Müller (2010*a*) proposes that edge features are inserted as long as the head H is still active; edge features are immediately discharged after their insertion, the remaining operation-inducing features on H are discharged afterwards. The former proposal rules out the patterns in (37-a) and (37-d) where a head H triggers operations *after* the edge feature has been discharged; the latter proposal is at odds with the pattern in (37-c) because the edge feature must have been inserted on the head H after all other operation-inducing features on H have been discharged, i.e., at a point where H was already inactive. Since these patterns are attested, we can conclude that either languages differ in the timing of edge feature insertion (some add edge features only to inactive heads and some only to active heads) or that we need a more flexible approach to edge feature insertion that allows for early as well as late insertion. I will not choose between these options and I will also not develop the second option here; I merely point out the relevance of the attested orderings in (37) to this.

Predicted and attested patterns:

	pattern (37-a)	pattern (37-b)	pattern (37-c)	pattern (37-d)
AAE	Trentino Fiorentino Ibibo (*both IM types bleed Agree*)	German English (*neither IM type bleeds Agree*)	Celtic Berber ... (*IM types have different effects*)	?
Interv.	Icelandic A Romance (*both IM types feed Agree*)	Icelandic C (*neither IM type feeds Agree*)	Icelandic B (*IM types have different effects*)	?
Poss	? (*both IM types bleed Agree*)	Hungarian (ϕ-Agree) (*neither IM type bleeds Agree*)	Hungarian (case) Nenets (ϕ-Agree) (*IM types have different effects*)	?

On variation in

(i) the AAE see Ouhalla (1993), Richards (1997), Ouali (2008), Phillips (2001) and references cited there.
(ii) intervention effects see Boeckx (1999), Anagnostopoulou (2003), Sigurðsson and Holmberg (2008).
(iii) possessor case/agreement see Nikolaeva (2002) and references cited there.

5.2. An Unattested Pattern (?)

The pattern in (37-d) that is predicted by the present analysis does not seem to be attested for any of the phenomena I have looked at. This might be accidental but I think it is striking that the same gap exists for three different empirical phenomena. What is more, it has been stated explicitly in the literature on the AAE that the reverse pattern has never been found, i.e., a language in which long Ā-movement of the subject bleeds subject-verb-agreement but short Ā-extraction does not (Ouhalla 1993).[13] Similarly, there exist three different Icelandic dialects, but apparently there is no dialect that exhibits the fourth pattern. I have never encountered an example of this pattern in the considerable literature on intervention effects. It remains to be seen whether this gap also exists for other phenomena. But given that it seems to be unattested for AAE and intervention, I assume that it does indeed reflect a property of the grammar and I will make a proposal as to what it might be.

I hypothesize that the absence of the fourth pattern is due to Specificity (for application of this concept to syntax see Sanders 1974, Pullum 1979, Lahne 2008*b*, van Koppen 2005) which states that more specific rules apply before less specific rules. With respect to operation-inducing features this means that the more specific IM-triggering feature is discharged before the less specific one. I suggest that non-edge features that trigger final movement steps are more specific than edge-features that trigger intermediate movement steps. The reason is that non-edge features attract elements of a certain category or with a certain effect on interpretation (e.g. [•D•], [•WH•], [•TOP•]), whereas edge features are categorially underspecified structure-building features that do not ask for a certain category or interpretative feature on the XP they attract. Hence, an order in which an edge feature is discharged before a non-edge feature, as is the case in the pattern in (37-d), is excluded by Specificity.[14]

[13] Ouhalla (1993: 486): "[. . .] there is apparently no language which displays the AAE in relation to long extraction but not in relation to short extraction. When there is an asymmetry, it seems to be invariably the case that the AAE holds in relation to short extraction but not in relation to long extraction."

[14] If this hypothesis is on the right track, it also predicts that external Merge to SpecH, which is standardly taken to be triggered by categorially fully specified features, must precede intermediate movement steps to SpecH. Note that this would be incompatible with the Intermediate Step Corollary presented in Müller (2010*b*). The interaction of Agree with external vs. internal Merge is beyond the scope of the present paper. I leave it for future reseach to determine whether this prediction is borne out.

This principle does not say anything about the order of probe vs. structure-building features. So their relative order is free.

5.3. Extrinsic vs. Intrinsic Ordering

A central question in the debate on rule ordering has always been whether orderings are extrinsic or intrinsic. An order is extrinsic if it must be stipulated in a language-specific fashion, i.e., two languages can differ solely in the order of rules. An order is intrinsic if it is determined by independent principles of the grammar (Chomsky 1965: 223, fn.6).[15] From a conceptual point of view, intrinsic orderings are to be preferred, but the question is whether all orderings are predicted by meta-principles of the grammar. Indeed, Pullum (1979) claims that all orderings are determined by Universal Principles, i.e., extrinsic ordering is not necessary. The major principles that Pullum argues to determine rule ordering in syntax (but also in other components of the grammar) are the following:

(i) Obligatory Precedence Principle (or Immediate Characterization, cf. Ringen 1972, Perlmutter and Soames 1979)
(ii) Specificity Principle (or Elsewhere Condition, cf. Sanders 1974, Anderson 1969 among many others)
(iii) Cyclic Principle (or the Cycle, cf. Chomsky et al. 1956, Chomsky and Halle 1968)

The Obligatory Precedence Principle states that obligatory rules apply before optional rules. This principle is not easily applicable in Minimalism because in this framework, operation-inducing features *must* be discharged if the context for the operation is met (cf. Pesetsky 1989, Pesetsky and Torrego 2001, Preminger 2011). It is only the presence of the feature on a head that can be optional, but if it is present, the operation it triggers must obligatorily apply. I will thus disregard this principle in what follows.

[15] Pullum (1979: ch.1) criticizes the definition of the terms 'extrinsic' and 'intrinsic' as introduced in Chomsky (1965). He thus replaces them by the terms 'parochial' for extrinsic and 'universal' for intrinsic orderings with slightly different defining properties, in particular concerning the question whether languages can differ solely in their order of rules. In what follows, I will stick to the original terminology because it is more widely used and, as Pullum acknowledges, because these terms have by and large been used in the way that he defines parochial and universal orderings, and I will do so as well.

The Specificity Principle demands that more specific operations apply before less specific operations. The consequences of this principle depend on what 'specific' means. It can, for example, refer to the context of a (transformational) rule: A rule R_1 is more specific than a rule R_2 if the context of R_1 includes the context of R_2. I have argued in section 5.2 that this principle may account for the absence of the order in (37-d).

Finally, the Cyclic Principle demands that rules in a lower cyclic domain apply before all rules in a higher cyclic domain:

(41) *Cyclic Principle*:
Any operation to the cyclic domain D_x will precede any operation to the cyclic domain D_{x-1}.

Traditionally, S and NP (= CP and DP in current terminology) are taken to constitute cyclic domains. Note that the Cyclic Principle does not have anything to say about the order of operations that apply within the same cycle. Since there are a lot of instances of rules that apply within a single clause and hence, within the same domain S, McCawley (1984, 1988) proposes to increase the number of cyclic domains and thereby to maximize the applicability as well as the predictive power of the Cyclic Principle. McCawley assumes that every constituent, (i.e. every projection in an incremental model of syntax) is a cyclic domain.[16] As he points out, it is very unlikely under this assumptions that two rules or operations apply within the same cycle. And indeed, he illustrates that many (if not all) of the orderings that needed to be stipulated under the original version of the Cyclic Principle are predicted by his stronger version of the principle. For applications of each of the three principles see Pullum (1979: ch.1).

With these principles in mind, we can now look at the variation in the order of probe features and the two types of IM-triggers summarized in (40). I claim that none of the principles introduced above can account for the full range of variation that is attested for the AAE, defective intervention effects and the encoding of possessors by case and/or agreement. The reason is that the patterns in (37-a) and (37-b) are the exact opposite of one another. It is impossible that

[16] A similar though not identical proposal is made in Williams (1974). If the Williams cycle is applied to the inclusion hierarchy that Williams adopts, this yields the same effects as McCawley's proposal as long as the rules or operations apply within the same clause; for cross-clausal interactions, the Williams cycle and McCawley's version of the Cyclic Principle make different predictions.

any principle predicts one order of features as well as the opposite. Note that this point is independent of the existence of the pattern in (37-d); even if only 3 out of the four predicted patterns are attested, no principle predicts all three of them.

Let me go into some more details. As indicated above, the Obligatory Precedence Principle is not applicable at all in a minimalist framework. As for the Specificity Principle, it may be able to distinguish between the two types of IM triggers in terms of specificity and thereby account for the absence of a pattern, but I do not see how it could distinguish between probe features and structure-building features in this respect. A probe and a structure-building feature (which is not an edge feature) do not seem to be of a different degree of specificity in any sense of this term. And even if one would find a property relative to which the two features have a different degree of specificity, the Specificity Principle would predict an order of the two features, but at the same time it would exclude the reverse order. As pointed out above, this is problematic given that both the patterns in (37-a) and (37-b) are attested.

The Cyclic Principle, which is the principle that has been said to be predict most of the rule orderings in syntax, does also not predict all of the attested patterns in (40). Obviously, in its original form with only S and NP nodes being cyclic domains, it cannot determine the orderings of operation-inducing features because all of the operations triggered by a single head H apply in the same S (CP) and NP (DP) cycle, respectively. Crucially, this also holds under McCawley's stricter version with every projection being a cyclic domain. Note that the cases in which a single head triggers more than one operation are exactly of the type that McCawley said would be rare, if not inexistent. Under one interpretation of McCawley's version of the principle, they do apply in the same projection. However, there are indeed two possible interpretations of McCawley's Cyclic Principle: Agree and Merge apply (i) either in the same cycle (H') or (ii) Agree applies in a lower cycle than IM (H′ vs. HP).[17] When looking at the point of the derivation where a head H that triggers Agree and IM is merged, the H′ projection is created and any of the operations that is triggered by H applies within this projection, i.e., they apply in the same cycle, cf. interpretation (i). However, one could also argue that Agree and IM apply in different cycles, cf. interpretation (ii): Since Agree applies under c-command,

[17] I thank Gereon Müller for pointing out to me this ambiguity in the interpretation of McCawley's version of the Cyclic Principle.

it will only affect the H′ constituent, whereas IM, when looking at the output of Merge, led to the projection of the HP and could thus be assumed to apply within HP. But under neither of these two interpretations does the stricter version of the Cyclic Principle predict the full range of attested variation: If interpretation (i) is assumed, the Cyclic Principle does not make any predictions because the operations apply in the same cyclic domain; thus, the principle is too weak. If one adopts interpretation (ii), Agree is wrongly predicted to always apply before IM, but the reverse order is also attested, cf. (40); in this second case, the principle is too strong. Thus, the Cyclic Principle also does not predict the full range of orders of operation-inducing features, not even under the stronger version with every projection being a cyclic domain.

The conclusion is that none of three principles allows for all the attested orderings. Some principles do not make any predictions at all, whereas others prohibit orders that are attested. Hence, extrinsic language-specific (parochial) ordering of the operation-inducing features is needed after all.

6. Conclusion

In this paper I have argued for a more fine-grained approach to the order of elementary operations: Merge not only needs to be ordered with respect to Agree; rather, one must distinguish between (at least) two different types of (internal) Merge because one type applies before and the other after Agree. Evidence for this split of internal Merge into two types comes from the observation that in some languages, if a head H triggers IM and Agree, IM has different consequences for Agree initiated by H, although on the surface the internally merged XP always occupies the same structural position SpecH. The behaviour of an internally merged XP with respect to its effect on the Agree relation can be predicted by the nature of the landing site: If SpecH is a final landing site for XP in a movement chain, XP movement feeds or bleeds Agree; if, however, SpecH is an intermediate landing site for XP, XP movement has the opposite effect, i.e., it counter-feeds or counter-bleeds Agree by H. If intermediate and final landing sites have different kinds of triggers, the opacity can be derived by a different ordering of the two IM triggers relative to Agree: Intermediate movement steps are triggered by edge features and apply after Agree; final movement steps are triggered by non-edge features and apply before Agree. The pattern that arises from this ordering is that edge-feature-driven IM of XP has the same effect on the Agree relation by H as if XP is not moved at all.

The analysis has the following general implications: (i) It presupposes a strictly derivational model of syntax. (ii) The variation predicted through reordering of operation-inducing features is indeed attested, except for one gap. I have proposed that this gap may be due to Specificity. (iii) The attested orderings provide evidence for the need of extrinsic ordering: None of the principles that have been proposed to determine the order of rules (or: operations) can predict the full range of variation. Most of the principles are too strong, disallowing certain attested orderings; the strict version of the Cyclic Principle proposed by McCawley (1988) is arguably too weak, making no predictions at all. (iv) Variation in defective intervention effects provides evidence that movement to SpecC goes through SpecT.

References

Abels, Klaus (2003): Successive Cyclicity, Anti-Locality, and Adposition Stranding. PhD thesis, University of Connecticut, Storrs, Connecticut.

Abels, Klaus (2004): Right Node Raising: Ellipsis or Across the Board Movement? *In:* K. Moulton and M. Wolf, eds, *Proceedings of NELS 34*. GLSA, Amherst, pp. 45–60.

Abels, Klaus (2012): *Phases: An Essay on Cyclicity in Syntax*. Mouton de Gruyter, Berlin.

Alexiadou, Artemis, Tibor Kiss and Gereon Müller (2012): *Local Modelling of Non-local Dependencies in Syntax*. Linguistische Arbeiten 547, de Gruyter, Berlin/New York.

Anagnostopoulou, Elena (2003): *The Syntax of Ditransitives: Evidence from Clitics*. Mouton de Gruyter, Berlin, New York.

Anderson, Stephen R. (1969): West Scandinavian Vowel Systems and the Ordering of Phonological Rules. PhD thesis, MIT, Cambridge, Massachusetts.

Anderson, Stephen R. (1974): *The Organization of Phonology*. Academic Press, New York.

Arregi, Karlos and Andrew Nevins (2012): *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Springer, Dordrecht.

Assmann, Anke, Svetlana Edygarova, Doreen Georgi, Timo Klein and Philipp Weisser (2013): Possessor Case in Udmurt: Multiple Case Assignment Feeds Post-syntactic Fusion. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar. Volume 90 of Linguistische Arbeitsberichte*. Institut für Linguistik: Universität Leipzig, pp. 23-63.

Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller and Philipp Weisser (2013): Ergatives Move too Early: On an Instance of Opacity in Syntax. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar. Volume 90 of Linguistische Arbeitsberichte*. Institut für Linguistik: Universität Leipzig, pp. 363-412.

Assmann, Anke and Fabian Heck (2013): Opaque Intervention. *In:* F. Heck and A. Assmann, eds, *Rule Interaction in Grammar. Volume 90 of Linguistische Arbeitsberichte*. Institut für Linguistik: Universität Leipzig, pp. 463-509.

Baker, Mark (2008): 'On the Nature of the Anti-Agreement Effect: Evidence from Wh-in-Situ in Ibibio', *Linguistic Inquiry* **39**, 615–632.

Baker, Mark and Nadya Vinokurova (2010): 'Two Modalities of Case Assignment: Case in Sakha', *Natural Language and Linguistic Theory* **28**, 593–642.

Baković, Eric (2011): Opacity and Ordering. *In:* J. Goldsmith, J. Riggle and A. Yu, eds, *The Handbook of Phonological Theory*. Wiley-Blackwell, Oxford, chapter 2.

Barss, Andrew (1986): Chains and Anaphoric Dependence. PhD thesis, MIT, Cambridge, Mass.

Bejár, Susana and Diane Massam (1999): 'Multiple Case Checking', *Syntax* **2**(2), 65–79.

Bobaljik, Jonathan (2008): Where's Phi? Agreement as a Post-Syntactic Operation. *In:* D. Harbour, D. Adger and S. Béjar, eds, *Phi-Theory: Phi Features Across Interfaces and Modules*. Oxford University Press, Oxford, pp. 295–328.

Boeckx, Cedric (1999): 'Conflicting C-Command Requirements', *Studia Linguistica* **53**(3), 227–250.

Brandi, Luciana and Patrizia Cordin (1998): Two Italian Dialects and the Null Subject Parameter. *In:* O. Jaeggli and K. Safir, eds, *The Null Subject Parameter*. Kluwer Academic Publishers, Dordrecht, pp. 111–142.

Brody, Michael (2002): On the Status of Representations and Derivations. *In:* S. Epstein and D. Seely, eds, *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford, pp. 19–41.

Béjar, Susana (2003): Phi-Syntax: A Theory of Agreement. PhD thesis, University of Toronto, Toronto.

Béjar, Susana and Milan Řezáč (2009): 'Cyclic Agree', *Linguistic Inquiry* **40**(1), 35–73.

Campos, Héctor (1997): 'On Subject Extraction and the Antiagreement Effect in Romance', *Linguistic Inquiry* **28**(1), 92–119.

Cheng, Lisa Lai-Shen (2006): 'Decomposing Bantu Relatives', *Proceedings of NELS* **36**, 197–215.

Chomsky, Noam (1951): Morphophonemics of Modern Hebrew. Master's thesis, University of Pennsylvania.

Chomsky, Noam (1957): *Syntactic Structures*. Mouton, The Hague and Paris.

Chomsky, Noam (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.

Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Mass.

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. *In:* M. Roger, M. David and J. Uriagereka, eds, *Step by Step*. MIT Press, Cambridge, Mass., pp. 89–155.

Chomsky, Noam (2001): Derivation by Phase. *In:* M. Kenstowicz, ed., *Ken Hale: A Life in Language*. MIT Press, Cambridge, Massachusetts, pp. 1–52.

Chomsky, Noam (2004): Beyond Explanatory Adequacy. *In:* A. Belletti, ed., *Structures and Beyond. The Cartography of Syntactic Structures Vol. 3*. Oxford University Press, Oxford.

Chomsky, Noam (2005): 'Three Factors in Language Design', *Linguistic Inquiry* **36**, 1–22.

Chomsky, Noam and Morris Halle (1968): *The Sound Pattern of English*. Harper & Row, New York.

Chomsky, Noam, Morris Halle and Fred Lukoff (1956): On Accent and Juncture in English. *In:* M. Halle, H. Lunt, H. McLean and C. van Schooneveld, eds, *For Roman Jakobson. Essays on the occasion of his sixtieth birthday*. Mouton, The Hague, pp. 65–80.

Chung, Sandra (1994): 'Wh-Agreement and 'Referentiality' in Chamorro', *Linguistic Inquiry* **25**, 1–44.

Clements, George, James McCloskey, Joan Maling and Annie Zaenen (1983): 'String-Vacuous Rule Application', *Linguistic Inquiry* **14**, 1–17.

Cole, Peter (1982): 'Subjacency and Successive Cyclicity: Evidence from Ancash Quechua', *Journal of Linguistic Research* **2**, 35–58.

Delsing, Lars-Olof (1998): Possession in Germanic. *In:* A. Alexiadou and C. Wilder, eds, *Possessors, Predicates and Movement in the Determiner Phrase*. Benjamins, Amsterdam, pp. 87–109.

Embick, David (2010): *Localism versus Globalism in Morphology and Phonology*. MIT Press, Cambridge, Mass.

Embick, David and Rolf Noyer (2001): 'Movement Operations after Syntax', *Linguistic Inquiry* **32**, 555–595.

Epstein, Samuel David and T. Daniel Seely (2002): Rule Applications as Cycles in a Level-Free Syntax. *In:* S. Epstein and T. Seely, eds, *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford, pp. 65–89.

Fox, Danny (2000): *Economy and Semantic Interpretation*. MIT Press, Cambridge, Mass.

Frampton, John (2004): Copies, Traces, Occurrences, and All That. Ms., Northeastern University.

Gallego, Ángel J. (2006): Phase Sliding. Ms., UAB/UMD.

Georgi, Doreen and Martin Salzmann (2011): 'DP-Internal Double Agreement is not

Double Agree: Consequences of Agree-Based Case Assignment within DP', *Lingua* **121**, 2069–2088.
Grohmann, Kleanthes K. (2003): 'Successive Cyclicity under (Anti-)Local Considerations', *Syntax* **6**, 260–312.
Halpert, Claire (2012): Argument Licensing and Agreement in Zulu. PhD thesis, MIT, Cambridge, Mass.
Heck, Fabian and Gereon Müller (2000): Successive Cyclicity, Long-Distance Superiority, and Local Optimization. *In:* R. Billerey and B. Lillehaugen, eds, *Proceedings of WCCFL*. Vol. 19, Cascadilla Press, Somerville, MA, pp. 218–231.
Heck, Fabian and Gereon Müller (2003): 'Derivational Optimization of Wh-Movement', *Linguistic Analysis* **33**, 97–148. (Volume appeared 2007).
Heck, Fabian and Gereon Müller (2007): Extremely Local Optimization. *In:* E. Brainbridge and B. Agbayani, eds, *Proceedings of the 26th WECOL*. California State University, Fresno, pp. 170–183.
Henderson, Brent (2009): 'Anti-Agreement and [person] in Bantu', *Selected Proceedings of the 38th Annual Conference on African Linguistics* **38**, 173–181.
Holmberg, Anders (2000): 'Scandinavian Stylistic Fronting: How any Category can Become an Expletive', *Linguistic Inquiry* **31**, 445–484.
Holmberg, Anders and Thorbjörg Hróarsdóttir (2003): 'Agreement and Movement in Icelandic Raising Constructions', *Lingua* **113**(10), 997–1019.
Jouitteau, Mélanie (2005): La syntaxe comparée du Breton. PhD thesis, University of Nantes, Nantes.
Kayne, Richard S. (1975): *French Syntax: The Transformational Cycle*. MIT Press, Cambridge, Mass.
Keine, Stefan (2010): *Case and Agreement from Fringe to Core: A Minimalist Approach*. Linguistische Arbeiten 536, de Gruyter, Berlin.
Kenstowicz, Michael and Charles Kisseberth (1977): *Topics in Phonological Theory*. Academic Press, NY.
Kenstowicz, Michael and Charles Kisseberth (1979): *Generative Phonology*. Academic Press, New York.
Kiparsky, Paul (1968): Linguistic Universals and Linguistic Change. *In:* R. Bach, E. Harms, ed., *Universals in Linguistic Theory*. Holt, New York, pp. 170–210.
Kiparsky, Paul (1971): Historical Linguistics. *In:* W. Dingwall, ed., *A Survey of Linguistic Science*. University of Maryland Linguistics Program, College Park, pp. 576–642.
Kiparsky, Paul (1976): Abstractness, Opacity, and Global Rules. *In:* A. Koutsoudas, ed., *The Ordering and Application of Rules*. Mouton.
Koppen, Marjo van (2005): One Probe – Two Goals: Aspects of Agreement in Dutch Dialects. PhD thesis, Leiden University., Leiden. LOT-publications number 105.

Koutsoudas, Andreas, Gerald Sanders and Craig Noll (1974): 'The Application of Phonological Rules', *Language* 50, 1–28.

Lahne, Antje (2008*a*): Excluding SVO in Ergative Languages: A New View on Mahajan's Generalisation. *In:* F. Heck, G. Müller and J. Trommer, eds, *Varieties of Competition*. Universität Leipzig, Institut für Linguistik, pp. 65–80. Vol. 87 of *Linguistische Arbeitsberichte*.

Lahne, Antje (2008*b*): Where There is Fire There is Smoke: Local Modelling of Successive-Cyclic Movement. PhD thesis, University of Leipzig.

Lasnik, Howard (2001): Derivation and Representation in Modern Transformational Syntax. *In:* M. Baltin and C. Collins, eds, *The Handbook of Contemporary Syntactic Theory*. Blackwell, Malden, Mass., pp. 62–88.

Lebeaux, David (1990): Relative Clauses, Licensing, and the Nature of the Derivation. *In:* J. Carter, R.-M. Déchaine, B. Philip and T. Sherer, eds, *Proceedings of NELS 20*. G.L.S.A., University of Massachusetts, Amherst, pp. 318–332.

Legate, Julie Ann (2002): Phases in "Beyond Explanatory Adequacy". Ms., MIT, Cambridge.

Marantz, Alec (1991): Case and Licensing. *In:* G. Westphal, B. Ao and H.-R. Chae, eds, *Proceedings of the Eighth Eastern States Conference on Linguistics*. University of Maryland, pp. 234–253.

McCawley, James D. (1984): Exploitation of the Cyclic Principle as a Research Strategy in Syntax. *In:* W. de Geest and Y. Putseys, eds, *Sentential Complementation*. Foris, Dordrecht, pp. 165–183.

McCawley, James D. (1988): *The Syntactic Phenomena of English*. Vol. 1, University of Chicago Press, Chicago.

McCloskey, James (1979): *Transformational Syntax and Model Theoretic Semantics. A Case Study in Modern Irish*. Reidel, Dordrecht.

McGinnis, Martha (2001): Variation in the Phase Structure of Applicatives. *In:* P. Pica and J. Rooryck, eds, *Linguistic Variation Yearbook 1*. John Benjamins, Amsterdam, pp. 101–142.

McGinnis, Martha J. (1998): Locality in A-Movement. PhD thesis, MIT, Cambridge, Mass.

Merchant, Jason (2006): Polyvalent Case, Geometric Hierarchies, and Split Ergativity. *In:* J. Bunting, S. Desai, R. Peachey, C. Straughn and Z. Tomkova, eds, *Proceedings of the 42nd Annual Meeting of the Chicago Linguistics Society*. Chicago Linguistics Society, Chicago, pp. 47–67.

Moravcsik, Edith A. (1995): Summing up Suffixaufnahme. *In:* F. Plank, ed., *Double Case: Agreement by Suffixaufnahme*. Oxford University Press, New York, pp. 451–484.

Müller, Gereon (2004): Phrase Impenetrability and Wh-Intervention. *In:* A.

Stepanov, G. Fanselow and R. Vogel, eds, *Minimality Effects in Syntax*. Mouton de Gruyter, Berlin, pp. 289–325.

Müller, Gereon (2009): Ergativity, Accusativity, and the Order of Merge and Agree. *In:* K. Grohmann, ed., *Explorations of Phase Theory. Features and Arguments*. Mouton de Gruyter, Berlin, pp. 269–308.

Müller, Gereon (2010*a*): 'On Deriving CED Effects from the PIC', *Linguistic Inquiry* 41, 35–82.

Müller, Gereon (2010*b*): Operator Islands, Maraudage, and the Intermediate Step Corollary. Ms., University of Leipzig. To appear in Balázs Surányi (ed.), *Minimalist Approaches to Syntactic Loality*.

Müller, Gereon (2011): *Constraints on Displacement: A Phase-Based Approach*. John Benjamins, Amsterdam.

Nikolaeva, Irina (2002): The Hungarian External Possessor in a European Perspective. *In:* C. Hasselblatt and R. Blokland, eds, *Finno-Ugrians and Indo-Europeans: Linguistic and Literary Contacts*. Shaker, Maastricht, pp. 272–285.

Ouali, Hamid (2008): On C-to-T-Feature Transfer: The Nature of Agreement and Anti-Agreement in Berber. *In:* R. D'Alessandro, S. Fischer and G. Hrafnbjargarson, eds, *Agreement Restrictions*. de Gruyter, pp. 159–180.

Ouhalla, Jamal (1993): 'Subject Extraction, Negation and the Anti-agreement Effect', *Natural Language and Linguistic Theory* 11, 477–518.

Perlmutter, David and Scott Soames (1979): *Syntactic Argumentation and the Structure of English*. The University of California Press, Berkeley.

Pesetsky, David (1989): 'Language Particular Processes and the Earliness Principle', Ms.

Pesetsky, David and Esther Torrego (2001): T to C Movement: Causes and Consequences. *In:* M. Kenstowicz, ed., *Ken Hale: A Life in Language*. MIT Press, Cambridge, Massachusetts, pp. 355–426.

Phillips, Colin (2001): Disagreement between Adults and Children. *In:* A. Mendikoetxea and M. Uribe-Etxebarria, eds, *Theoretical Issues on the Morphology-Syntax Interface*. ASJU, San Sebastian, pp. 359–394.

Preminger, Omer (2011): Agreement as a Fallible Operation. PhD thesis, Massachusetts Institute of Technology, Cambridge, Mass.

Pullum, Geoffrey K. (1979): *Rule Interaction and the Organization of a Grammar*. Garland, New York & London. (PhD thesis from 1976, University of London).

Řezač, Milan (2004): Elements of Cyclic Agree: Agree and Merge. PhD thesis, University of Toronto.

Richards, Marc D. (2007): 'On Feature-Inheritance: an Argument from the Phase Impenetrability Condition', *Linguistic Inquiry* 38, 563–572.

Richards, Norvin (1997): What Moves Where When in Which Language? PhD thesis, MIT, Cambridge, Mass.

Ringen, Catherine (1972): 'On Arguments for Rule Ordering', *Foundations of Language* **8**, 266–273.

Rizzi, Luigi (2004): On the Form of Chains: Criterial Positions and ECP Effects. Ms., University of Siena.

Rizzi, Luigi (2007): 'On Some Properties of Criterial Freezing', *CISCL Working Papers on Language and Cognition. STIL* **1**.

Rizzi, Luigi and Ur Shlonsky (2007): Strategies for Subject Extraction. *In:* U. Sauerland and H.-M. Gärtner, eds, *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*. Mouton de Gruyter, Berlin, pp. 115–160.

Ross, John R. (1967): Constraints on Variables in Syntax. PhD thesis, MIT, Cambridge, Mass.

Sanders, Gerald (1974): 'Precedence Relations in Language', *Foundations of Language* **11**, 361–400.

Schneider-Zioga, Patricia (2007): 'Anti-Agreement, Anti-Locality and Minimality. The Syntax of Dislocated Subjects', *Natural Language and Linguistic Theory* **25**, 403–446.

Sigurðsson, Halldór Ámann and Anders Holmberg (2008): Icelandic Dative Intervention: Person and Number are Separate Probes. *In:* R. d'Alessandro, G. H. Hrafnbjargarson and S. Fischer, eds, *Agreement Restrictions*. Mouton de Gruyter, Berlin, pp. 251–280.

Sproat, Richard (1985): 'Welsh Syntax and VSO Structure', *Natural Language and Linguistic Theory* **3**, 173–216.

Starke, Michal (2001): Move Dissolves Into Merge: A Theory of Locality. PhD thesis, University of Geneva.

Sternefeld, Wolfgang (2006): *Syntax*. Stauffenburg, Tübingen. Two volumes.

Szabolcsi, Anna (1994): The Noun Phrase. *In:* F. Kiefer, ed., *Syntax and Semantics, vol. 27: The Structure of Hungarian*. Academic Press, San Diego, pp. 179–274.

Torrego, Esther (1984): 'On Inversion in Spanish and Some of Its Effects', *Linguistic Inquiry* **15**, 103–129.

Unger, Christina (2010): A Computational Approach to the Syntax of Displacement and the Semantics of Scope. PhD thesis, Universiteit Utrecht.

Vries, Marc de (2006): 'Possessive Relatives and (Heavy) Pied Piping', *The Journal of Comparative Germanic Linguistics* **9**, 1–52.

Williams, Edwin S. (1974): Rule Ordering in Syntax. PhD thesis, MIT, Cambridge, Mass.